C2

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 September 2001 (27.09.2001)

PCT

(10) International Publication Number
WO 01/70982 A2

(51) International Patent Classification[7]: C12N 15/12, 15/11, C07K 14/47, A61K 31/713, C12Q 1/68, G01N 33/68 // C12N 9/90, A61P 35/00

(21) International Application Number: PCT/US01/09559

(22) International Filing Date: 23 March 2001 (23.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/536,058 23 March 2000 (23.03.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 09/536,058 (CIP)
Filed on 23 March 2000 (23.03.2000)

(71) Applicant *(for all designated States except US)*: **IMMUSOL INCORPORATED** [US/US]; 10790 Roselle Street, San Diego, CA 92121 (US).

(71) Applicant and
(72) Inventor: **BEGER, Carmela** [DE/DE]; Kronsberger Strasse 20, 30559 Hanover (DE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **BARBER, Jack** [US/US]; 11168 Charlota Street, San Diego, CA 92129 (US). **WONG-STAAL, Flossie** [US/US]; 12737 Monterey Cypress Way, San Diego, CA 92130 (US).

(74) Agents: **REITER, Stephen, E.** et al.; Foley & Lardner, 23rd Floor, 402 W. Broadway, San Diego, CA 92101-3542 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/70982 A2

(54) Title: BRCA-1 REGULATORS AND METHODS OF USE

(57) **Abstract:** The invention provides ribozymes and encoding nucleic acids having target recognition sequences that allow the ribozyme to target and cleave BRCA-1 regulators, resulting in upregulation of BRCA-1 in a cell. Also provided are nucleic acids encoding BRCA-1 regulators that contain the target sequences recognized by the ribozymes of the invention. Fragments of these nucleic acid and protein sequences also are provided. Further provided is a method for identifying a gene where the expression level is affected by a BRCA-1 regulator and the identity of several such affected genes. Still further provided is a method of identifying a compound that modulates the activity of a BRCA-1 regulator. Also provided is a method of treating cancer, comprising introducing a ribozyme selectively reactive with an RNA encoding a BRCA-1 regulator into a cancerous cell. The invention further comprises a method of detecting a neoplastic cell in a sample.

BRCA-1 REGULATORS AND METHODS OF USE

FIELD OF THE INVENTION

This invention relates generally to proliferative diseases such as cancer and, more specifically, to regulation of the tumor suppressor BRCA-1 that can be used to diagnose and treat proliferative diseases.

BACKGROUND OF THE INVENTION

Cancer is one of the leading causes of death in the United States. Each year, more than half a million Americans die from cancer, and more than one million are newly diagnosed with the disease. Cancerous tumors result when a cell escapes from its normal growth regulatory mechanisms and proliferates in an uncontrolled fashion. Tumor cells can metastasize to secondary sites if treatment of the primary tumor is either not complete or not initiated before substantial progression of the disease. Early diagnosis and effective treatment of tumors is, therefore, essential for survival.

Cancer involves the clonal replication of populations of cells that have gained competitive advantage over normal cells through the alteration of regulatory genes. Regulatory genes can be broadly classified into "oncogenes" which, when activated or overexpressed, promote unregulated cell proliferation, and "tumor suppressor genes," which when inactivated or underexpressed, fail to prevent abnormal cell proliferation. Loss of function or inactivation of tumor suppressor genes is thought to play a central role in the initiation and progression of a significant number of human cancers.

Mutations in one known tumor suppressor gene, BRCA-1, contribute in essentially all cases to inherited susceptibility to ovarian and breast cancers. Additionally, BRCA-1 expression levels are reduced or undetectable in the tumor cells of sporadic breast cancers.

Approaches for treating cancer by modulating the function of tumor suppressor genes, either with pharmaceutical compounds or by gene therapy methods, have yielded promising results in animal models and in human clinical trials. Approaches for diagnosing and prognosing cancer by identifying mutations in known tumor suppressor genes or mutations in genes that regulate the expression of

tumor suppressor genes have also been developed. For example, identifying individuals containing germline mutations in known tumor suppressor genes has permitted the identification of individuals at increased risk of developing cancer. Such individuals are then closely monitored or treated prophylactically to improve their chance of survival. Identifying the pattern of alterations of known tumor suppressor genes in biopsy samples is also being used to determine the presence or stage of a tumor. Being able to determine whether a cancer is benign or malignant, at an early or late stage of progression, provides the patient and clinician with a more accurate prognosis and can be used to determine the most effective treatment for the patient.

In view of the importance of tumor suppressor molecules in the detection and treatment of cancer, and the known correlation of the tumor suppressor BRCA-1 with breast and ovarian cancers, there exists a need to identify nucleic acids and polypeptides that influence the level or activity of BRCA-1. The present invention satisfies this need and provides related advantages as well.

SUMMARY OF THE INVENTION

The invention is directed to methods of identifying genes that contribute to the formation of cancer and to the identified genes and gene products as therapeutic targets for the treatment of cancer. The methods are directed to the identification of genes encoding proteins that regulate the tumor suppressor, BRCA-1, resulting in an increased propensity for cancer. Accordingly, the present invention provides ribozymes and encoding nucleic acids having target recognition sequences that enable the ribozymes to bind and cleave BRCA-1 regulators, resulting in upregulation or downregulation of BRCA-1 in a cell.

Also provided are nucleic acids encoding BRCA-1 regulators that contain the target sequences recognized by the ribozymes of the invention. Fragments of these nucleic acid and protein sequences also are provided.

Further provided is a method for identifying a gene for which the expression level is affected by a BRCA-1 regulator comprising: a) hybridizing a first mRNA and a second mRNA to at least one candidate gene, wherein the first mRNA is obtained from cells expressing a ribozyme that targets and cleaves mRNA encoding

a BRCA-1 regulator and wherein the second mRNA is from control cells otherwise similar to those expressing the ribozyme except that the BRCA-1 regulator mRNA is not targeted by a ribozyme; and b) comparing the relative amounts of the first and second mRNA that hybridizes to the gene. The mRNAs may be reverse transcribed into DNA before hybridization and the gene or the mRNAs (or cDNA copies) may be labeled with a detectable moiety such as a fluorescent dye. In a further embodiment, hybridization to multiple genes is achieved by arraying the genes on a solid support.

Yet further provided is a method of identifying a compound that modulates the activity of a BRCA-1 regulator, the method comprising contacting a BRCA-1 regulator with a test compound and a target molecule responsive to the activity of the BRCA-1 regulator, wherein an increase or decrease in the activity of the BRCA-1 regulator for the target molecule in the presence of the test compound as compared to the absence of the test compound identifies a compound that modulates the activity of a BRCA-1 regulator. The activity of the BRCA-1 regulator can be a DNA binding activity, protein kinase activity, GTP binding activity, protease activity, or protein binding activity.

Still further provided is a method of treating cancer, comprising contacting a cancer cell with an expression vector that encodes a ribozyme selectively reactive with an RNA encoding a BRCA-1 regulator or contacting the cancer cell with the ribozyme.

Also provided is a method of detecting a neoplastic cell in a sample wherein the neoplastic cell is associated with an altered expression of a BRCA-1 regulator or an altered structure of a BRCA-1 regulator as compared to a normal cell, comprising: (a) contacting the sample with a detectable agent specific for a BRCA-1 regulator nucleic acid or an encoded BRCA-1 regulator polypeptide; and (b) detecting the nucleic acid or polypeptide in the sample, wherein altered expression or structure of the polypeptide indicates the presence of a neoplastic cell in the sample.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a haripin ribozyme and a substrate RNA. (SEQ ID NOS: 3 and 99).

Figure 2 shows alignments of the (a) nucleic acid; and (b) amino acid sequences for BR1 (SEQ ID NOS: 1 and 2, respectively) and Thema EFG (SEQ ID NOS: 96 and 97, respectively).

Figure 3 shows rounds of cell sorting in selection for cells with increased EGFP expression.

Figure 4 shows the enrichment of EGFP expressing cells after rounds of cell sorting.

Figure 5 shows a RNA expression levels after a fifth round of cell sorting.

Figure 6 shows a schematic of a method for amplifying DNA corresponding to the sequence of a TST.

Figure 7 shows an RST for BR1 and sites of binding for several RSTs of BR1 (SEQ ID NOS: 100 and 101, respectively).

Figure 8 shows validation ribozyme sequences (SEQ ID NOS: 102 through 221).

Figure 9 show the nucleic acid sequence of BR1 (SEQ ID NO: 1) and predicted translated polypeptide sequence (SEQ ID NOS: 2 and 222 through 231).

DETAILED DESCRIPTION OF THE INVENTION

The invention is directed to methods of identifying genes that contribute to the formation of cancer and to the identified genes and gene products as therapeutic targets for the treatment of cancer. The methods are directed to the identification of genes encoding proteins that regulate the tumor suppressor, BRCA-1, resulting in an increased propensity for cancer. Genes encoding BRCA-1 regulators are sought for identification and as therapeutic targets because decrease in the level of BRCA-l expression or BRCA-1 activity is linked with certain cancers, for example, breast carcinomas.

In one embodiment, BRCA-1 regulators have been identified. BRCA-l is a tumor suppressor which prevents cells from undergoing uncontrolled growth such as occurs with cancer cells. BRCA-l can act as a transcription factor, and thus BRCA-1

tumor suppressor activity can result, at least in part, from regulation of gene expression. Decreased BRCA-1 activity is associated with cancers such as breast cancer, ovarian cancer, endometrial cancer, prostate cancer, and the like. While in some cancers, the decreased BRCA-1 activity is due to mutations in the BRCA-1 gene, in other cancers, decreased BRCA-1 activity is due to reduced levels of BRCA-1 expression. To date, regulation of BRCA-1 expression has not been well characterized. In the present invention, numerous nucleic acid sequences have been identified which correspond to BRCA-1 regulators. These nucleic acid sequences have been shown to regulate BRCA-1 gene expression or activity. Ribozymes that inhibit expression of many of these nucleic acid sequences, including BBC1, BR1, ID4, BR2, and BR3, can increase BRCA-1 expression or activity.

As used herein, the term "substantially pure"' when used in reference to .a nucleic acid or polypeptide of the invention is intended to mean a molecule that is in a form that is relatively free from cellular components such as lipids, polypeptides, nucleic acids or other cellular material that it is associated with in its natural state.

As used herein, the term "nucleic acid" is intended to mean a single- or double-stranded DNA or RNA molecule. For example, a nucleotide designated as "T" also is equivalent to a "U" nucleotide in a recited sequence. A nucleic acid molecule of the invention can be of linear, circular or branched configuration, and can represent either the sense or antisense strand, or both, of a native nucleic acid molecule. Unless otherwise indicated, a reference to a nucleotide sequence of a nucleic acid molecule includes the sequence in single stranded form and in double stranded form. The term also is intended to include nucleic acid molecules of both synthetic and natural origin. A nucleic acid molecule of natural origin can be derived from any animal, such as a human, non-human primate, mouse, rat, rabbit, bovine, porcine, ovine, canine, feline, or amphibian, or from a lower eukaryote, such as Drosophila, C. elegans or yeast. A synthetic nucleic acid includes, for example, nucleic acids prepared by chemical and enzymatic synthesis. The term "'nucleic acid" is similarly intended to include analogues of natural nucleotides which have similar binding properties as the referenced nucleic acid and which can be utilized in a manner similar to naturally occurring nucleotides and nucleosides.

As used herein, the term "fragment" when used in reference to a nucleic acid is intended to mean a portion or segment of the nucleic acid molecule having the ability to selectively hybridize or bind to the subject nucleic acid, or its complement. The term "selectively hybridize" as used herein, refers to the ability of a nucleic acid or fragment to bind the subject nucleic acid molecule without substantial cross-reactivity with a molecule that is not the subject nucleic acid molecule.

A fragment of a nucleic acid molecule of the invention includes at least about 8-12 nucleotides of the subject nucleic acid. For example, a ribozyme sequence tag (RST) of a hairpin ribozyme and its complementary target sequence tag (TST) described herein generally include about 14-16 nucleotides of which 6-8 nucleotides are in helix 1, about 4 nucleotides are in helix 2, the two helixes separated by 4 non-base pairing nucleotides. Not all of the RST nucleotides in helix 1 need be complementary to the TST for the ribozyme to target and cleave the TST containing RNA. Generally, 1-2 bases of RST helix 1, depending on the position of the nucleotide in the sequence, may be non-complementary to the TST without loss of ribozyme function, although in some cases, up to 4 of bases may be non-complementary. Therefore, a fragment having the ability to selectively hybridize can contain about 8, 9, 10, 11 or 12 nucleotides of the subject nucleic acid. A fragment can also contain a greater number of nucleotides corresponding to the subject nucleic acid, or complement thereof, including for example, about 13, 14 or 15 nucleotides as well as at least 16, 17, 18, 19 or 20 nucleotides so long as it maintains the ability to selectively hybridize to the subject nucleic acid. Additionally, a fragment can be longer, including at least about 25, 30, 40, 50, 100, 300 or 500 or more nucleotides, and can include up to the full length of the reference nucleic acid molecule minus one nucleotide. Fragments of such lengths are able to selectively hybridize with the subject nucleic acid molecule in a variety of detection formats described herein and known to those skilled in the art.

Therefore, a fragment of a nucleic acid molecule of the invention can be used, for example, as an RST to target a ribozyme to a nucleic acid of the invention or as a selective inhibitor of uncontrolled growth in a cell having abnormal BRCA-1 activity or expression levels. A TST can be used as a PCR primer to selectively amplify a nucleic acid molecule of-the invention; as a selective primer for 5' or 3'

RACE to determine or identify 5' or 3' sequence of a nucleic acid molecule identified in methods of the invention; as a selective probe to identify or isolate a nucleic acid molecule of the invention on a RNA or DNA blot, or from a genomic or cDNA library.

The term "unique" when used in reference to a specific nucleic acid fragment is intended to mean a fragment of the subject nucleic acid that contains at least one nucleotide at a particular position that is characteristic, distinct or novel when compared to a different nucleotide sequence, or a related nucleotide sequence at the same or analogous position. In reference to a particular sequence, for example, the nucleotide sequence encoding the BRCA-1 regulator 1, or BR1 (SEQ ID NO: 1) differs from the elongation factor G (EFG) encoding sequence at about 285 codon positions or about 694 nucleotides within the coding region. Therefore, for each of these codon positions, there is at least one nucleotide which differs from the EFG sequence at that position and is therefore characteristic of the BR1 nucleotide sequence. A BR1 nucleic acid fragment containing one such characteristic nucleotide is a unique fragment.

As used herein, the term "substantially the same" when used in reference to a nucleotide sequence is intended to mean a nucleic acid molecule that retains its ability to selectively hybridize to the reference nucleic acid. Therefore, a nucleic acid molecule having substantially the same sequence compared to a reference nucleic acid can include, for example, one or more additions, deletions or substitutions with respect to the reference sequence so long as it can selectively bind to that sequence. Included within this definition are encoding nucleic acids that have degenerate codon sequences at one or more positions and therefore differ in nucleotide sequence compared to the reference nucleic acid but substantially maintain the referenced encoded amino acid sequence.

As used herein the term "substantially the same," when used in reference to a polypeptide of the invention, is intended to mean an amino acid sequence that contains minor modifications with respect to the reference amino acid sequence, so long as the polypeptide retains one or more of the functional activities exhibited by the polypeptide as a whole. A polypeptide that has substantially the same amino acid sequence as a reference human amino acid sequence can be, for example, a

homologous polypeptide from a vertebrate species, such as a non-human primate, mouse, rat, rabbit, bovine, porcine, ovine, canine, feline, or amphibian, or from a lower eukaryote, such as Drosophila, C. elegans or yeast.

A polypeptide that has substantially the same amino acid sequence as a reference sequence can also have one or more deliberately introduced modifications, such as additions, deletions or substitutions of natural or non-natural amino acids, with respect to the reference sequence. Those skilled in the art can determine appropriate modifications that, for example, serve to increase the stability, bioavailability, bioactivity or immungenicity of the polypeptide, or facilitate its purification, without altering the desired functional activity. For example, introduction of a D-amino acid or an amino acid analog, or deletion of a lysine residue, can stabilize a polypeptide and reduce degradation. Likewise, addition of tag sequences, such as epitope or histidine tags, or sorting sequences, can facilitate purification of the recombinant polypeptide. Depending on the modification and the source of the polypeptide, the modification can be introduced into the polypeptide, or into the encoding nucleic acid sequence.

Computer programs known in the art, for example, DNASTAR software, can-be used to determine which amino acid residues can be modified as indicated above without abolishing the desired functional activity. Additionally, guidance in modifying amino acid sequences while retaining functional activity is provided by aligning homologous BRCA-1 regulator polypeptides from various species. Those skilled in the art understand that evolutionarily conserved amino acid residues and domains are more likely to play a role in the biological activity than less well-conserved residues and domains.

In general, an amino acid sequence that is substantially the same as a reference amino acid sequence will have greater than about 25% identity with the reference sequence, although an amino acid sequence that is substantially the same as a reference sequence can have greater than about 30% identity, greater than about 40% identity, greater than about 50% identity, greater than about 60% identity, greater than about 70% identity, preferably greater than about 80% identity, more preferably greater than about 90% identity, and especially preferably greater than about 95% identity, or greater than about 98% identity. The amino acid sequences

which align across two sequences, and the presence of gaps and non-homologous regions in the alignment, can be determined by those skilled in the art based, for example, on a BLAST 2 or Clustal V or similar computer alignment software. A computer alignment can, if desired, be optimized visually by those skilled in the art. The percent identity of two sequences is determined as the percentage of the total amino acids that align in such an alignment which are identical. Those skilled in the art understand that two amino acid molecules with a given percentage identity over the entire sequence or over a substantial portion or portions thereof, are more likely to exhibit similar functional activities than two molecules with the same percentage identity over a shorter portion of the sequence. Sequence identity is preferably determined with by BLAST searching with the default settings provided at the website of the National Cancer Biological Institute (NCBI).

As used herein, the term "functional fragment" when used in reference to a polypeptide of the invention is intended to refer to a portion, segment or fragment of the polypeptide, which retains at least one of the activities of the full length polypeptide. For example, a functional fragment of any of the BRCA-l regulators of the invention can be a portion of the polypeptide that maintains its ability to regulate BRCA-1 expression or activity. For the specific example of ID4, a functional fragment of ID4 can be a portion of ID4 that maintains its ability to bind with one or more helix-loop-helix transcription factors or a portion of ID4 that modulates differentiation in adipocytes or neuronal cells.

As used herein, the term "ribozyme sequence tag" or "RST" is intended to mean the target recognition domain of a ribozyme. Therefore, the structure of an RST hairpin ribozyme can be 5'-N_8-AGAA-N_4-3' where N_8 and N_4 are complementary to sequences of the target RNA . The bases AGAA form a non-binding loop with the NGUC sequence of the target RNA. Therefore, a "target sequence tag" nucleic acid or "TST" as used herein, is a nucleic acid having a nucleotide sequence that is capable of selectively hybridizing to an RST of a ribozyme and being cleaved by the ribozyme. For example, the TST regions capable of selectively hybridizing to the RST will be substantially the complement of the helix 1 and helix 2 RST region sequences. These selectively hybridizing regions are separated by, for example, a GUC which is capable of being cleaved by a hairpin

ribozyme and therefore will have the structure 5'-N_5-GUC-N_8-3' where the first four nucleotides of N_5 represent the TST complementary sequence of the ribozyme helix 2 and N_8 represents the TST complementary sequence of the ribozyme helix 1.

As used herein, the term "ribozyme" or "ribozyme RNA molecule" is intended to mean a catalytic RNA that cleaves RNA. Ribozymes include both hairpin and hammerhead classes, which differ in mechanism for hybridization. The term "hairpin ribozyme" is intended to refer to an RNA molecule having the general nucleic acid sequence and two-dimensional configuration of the molecule shown in Figure 1, which is capable of selectively hybridizing, or of both selectively hybridizing and cleaving, a target RNA. The term is also intended to include both hairpin ribozyme RNA molecules as well as single- and double-stranded DNA molecules that, when expressed, form hairpin ribozyme RNA molecules. Generally, a hairpin ribozyme will have from about 50 to 54 nucleotides, which form two helical domains (helix 3 and helix 4) and 3 loops (Loops 2, 3 and 4). Two additional helices, helix 1 and helix 2, form between the ribozyme and its RNA target or substrate (Figure 1). A hairpin ribozyme binds a target RNA by forming Watson-Crick base pairs between the substrate and helix 1 and helix 2 sequences (see dots in Figure 1; "N" is any nucleotide). The length of helix 2 is usually about 4 nucleotides, and the length of helix 1 can vary from about 6-10 nucleotides or more. A hairpin ribozyme can have catalytic activity, and thus cleave the target RNA such as at the cleavage site shown in Figure 1. The catalytic activity of the hairpin ribozyme also can be disabled by, for example, altering the AAA sequence in Loop 2 to CGU. Those skilled in the art can determine which modifications to the overall hairpin ribozyme structure can be made and still maintain the target binding, or both target binding and catalytic activity of a hairpin ribozyme of the invention.

As used herein, the term "library" or "ribozyme library" is intended to mean a collection or population of different species of ribozyme RNA molecules. Within a population, any of the ribozyme species can be uniquely represented or redundant. Therefore, the term "randomized" or "random" when used in reference to a ribozyme library is intended to refer to a population of ribozymes that have differing nucleotide sequences in their target recognition sequence. The differing nucleotide sequences can be purposefully introduced, such as by degenerate, partially

degenerate or variegated oligonucleotide synthesis, or other methods well known to those skilled in the art. Alternatively, the differing nucleotide sequences can be introduced by a variety of mutagenesis methods including, for example, chemical and enzymatic methods well known in the art. A random ribozyme library also can be assembled, for example, from combining a collection of different ribozyme species into a single population. Details on the synthesis and construction of random ribozyme libraries is described in the Examples and in WO 00/05415 to Barber et al.

As used herein, the term "target recognition sequence" when used in reference to a ribozyme is intended to mean the substrate binding site of a ribozyme, which corresponds to an RST RNA nucleotide sequence or which corresponds to the complement of an TST nucleic acid nucleotide sequence. For the specific example of a hairpin ribozyme, the target recognition sequence corresponds to the nucleotide sequences of the helix 1 or helix 2 domain or both (see Figure 1). The target recognition sequences of helix 1 and 2 can be separated by catalytic nucleotides, which in the specific example of a hairpin ribozyme correspond to the nucleotides AGAA.

As used herein, the term "BRCA-1 regulator" is intended to mean a gene product or a nucleic acid, such as a structural or functional RNA, that modulates the expression or activity of BRCA-1. A BRCA-1 regulator can modulate BRCA-1 expression through, for example, direct binding of a BRCA-1 regulator to a nucleic acid element that influences BRCA-1 expression, such as the BRCA-1 5' regulatory region, where the BRCA-1 regulator bound to the nucleic acid element results in an increase or decrease in BRCA-1 expression. Another manner in which a BRCA-1 regulator can modulate BRCA-1 expression is by binding to a separate molecule that modulates BRCA-1 expression, such as a transcription factor, in such a way that results in an increase or decrease in BRCA-1 expression. BRCA-1 expression can also be modulated by a BRCA-1 regulator which increases or decreases the translation of BRCA-1-encoding mRNA. Additionally, a BRCA-1 regulator can bind to BRCA-1, thereby modulating BRCA-1 activity. For example, a BRCA-1 regulator can bind BRCA-1 thereby increasing or decreasing the transcriptional activation activity of BRCA-1. Similarly, BRCA-1 regulator can modulate BRCA-1 activity by binding a separate molecule that modulates BRCA-1 expression, for

example a transcription factor such as TFIIH, and thereby increase or decrease BRCA-1 activity. It is not necessary that the BRCA-1 regulator be essential for modulation of BRCA-l expression or activity. All that is needed is that the BRCA-1 regulator be functionally involved in modulating BRCA-1 expression or activity. A BRCA-1 regulator can be encoded by a gene including, for example, a gene originating from the species of the infected cell type as well as a heterologous gene that becomes incorporated into the cell's genome. Also, in accordance with the invention, regulators of BRCA-1 can be involved in regulating BRCA-2 or other factors involved in the DNA repair response.

As used herein, the term "modulate" when used in reference to BRCA-1 expression or activity refers to the ability of a BRCA-1 regulator to alter BRCA-1 expression or activity. Altered expression or activity includes, for example, an increase or decrease in BRCA-1 expression or activity.

As used herein, the term "BRCA-1" is intended to mean a member of the BRCA-1 tumor suppressor family. Exemplary members of the tumor suppressor family are human BRCA-1 and species homologs of BRCA-1.

As used herein, the term "selection marker" means a natural or modified gene product that influences cell viability or cell growth, or can be used as a measurable indicator of an activity associated with a cell such as viability, growth, gene expression and the like. Specific examples include enhanced green fluorescence protein (EGFP), hygromycin resistance gene and the tumor suppressor BRCA-1. For example, in the presence of hygromycin, the expression of the hygromycin resistance gene permits cells to survive. In a contrasting example, expression of BRCA-1 suppresses anchorage-independent cell growth, resulting in, for example, decreased ability to grow in soft agar. EGFP is another selection marker that can be used as a measurable indicator of gene expression. Use of EGFP allows a specific, user-defined expression level to be selected for fluorescence activated cell sorting (FACS) and other cell analysis methods. Iterative FACS-based selection results in enrichment of cells expressing EGFP at the user-defined expression level.

As used herein, the term "treating" when used in reference to cancer is intended to mean a reduction in the severity, or prevention of a neoplastic disease. Therefore, "'treating a cancerous disease" as used herein, is intended to mean a

reduction in severity, regression, or prevention of a cancerous disease. Reduction in severity includes, for example, an arrest or a decrease in clinical symptoms, physiological indicators or biochemical markers. Prevention of the cancer includes, for example, precluding the occurrence of the cancer, such as in prophylactic uses to individuals having a known or suspected propensity for developing cancer, or reversing the cancer in an individual to their non-diseased state of health.

As used herein, the terms "cancer," "tumor" and "neoplasm" and grammatical variants thereof are intended to mean abnormal growth in a tissue or organ. Typically, such a growth is characterized by uncontrolled cell proliferation and can be malignant or benign.

As used herein, the term "neoplastic cell" is intended to mean a cell that has altered expression or activity of a BRCA-1 regulator compared to a normal cell from the same or a different individual. A neoplastic cell will generally also exhibit histological or proliferative features of a malignant or premalignant cell. For example, histological methods can be used to show invasion of surrounding normal tissue, increased mitotic index, increased nuclear to cytoplasmic ratio, altered deposition of extracellular matrix, or a less differentiated cell phenotype. A neoplastic cell can also exhibit unregulated proliferation, such as anchorage independent cell growth, proliferation in reduced-serum medium, loss of contact inhibition, or rapid proliferation compared to normal cells.

As used herein, the term "altered expression" of a BRCA-1 regulator nucleic acid detected by a method of the invention refers to an increased or decreased amount of a BRCA-1 regulator nucleic acid in the test sample relative to known levels in a normal sample. Altered abundance of a nucleic acid molecule can result, for example, from an altered rate of transcription, from altered transcript stability, or from altered copy number of the corresponding gene.

As used herein, the term "altered structure" of a nucleic acid molecule refers to differences, such as point mutations, deletions, translocations, splice variations and other rearrangements, between the structure of a BRCA-1 regulator nucleic acid molecule in a test sample and the structure of the BRCA-1 regulator nucleic acid molecule in a normal sample. Those skilled in the art understand that mutations that alter the structure of a nucleic acid molecule can also alter its expression.

As used herein, the term "altered expression" of a BRCA-1 regulator polypeptide refers to an increased or decreased amount, or altered subcellular localization of the polypeptide in the test sample relative to known levels or localization in a normal sample. Altered abundance of a polypeptide can result, for example, from an altered rate of translation or altered copy number of the corresponding message, or from altered stability of the protein. Altered subcellular localization can result, for example, from truncation or inactivation of a sorting sequence, from fusion with another polypeptide sequence, or from altered interaction with other cellular polypeptides.

As used herein, the term "altered structure" of a polypeptide refers to differences in amino acid sequence, post-translational modifications, or conformation, of the polypeptide in the test sample relative to a normal sample which result in altered expression or activity of a BRCA-1 regulator polypeptide. Post-translational modifications include, for example, phosphorylation, glycosylation and acylation. Such differences can be detected, for example, with a structure-specific detectable binding agent.

As used herein, the term "sample" is intended to mean any biological fluid, cell, tissue, organ or portion thereof, that includes or potentially includes nucleic acids and polypeptides of the invention. The term includes samples present in an individual as well as samples obtained or derived from the individual. For example, a sample can be a histologic section of a specimen obtained by biopsy, or cells that are placed in or adapted to tissue culture. A sample further can be a subcellular fraction or extract, or a crude or substantially pure nucleic acid or protein preparation. A sample can be prepared by methods known in the art suitable for the particular format of the detection method.

As used herein, the term "detectable agent" refers to a molecule that renders a BRCA-1 regulator nucleic acid or polypeptide detectable by an analytical method. An appropriate detectable agent depends on the particular detection format, and can be determined for a particular application of the method by those skilled in the art. For example, a detectable agent specific for a BRCA-1 regulator nucleic acid molecule can be a complementary nucleic acid molecule, such as a hybridization probe or non-catalytic ribozyme, that selectively hybridizes to the nucleic acid

molecule. A hybridization probe or ribozyme can be labeled with a detectable moiety, such as a radioisotope, fluorochrome, chemiluminescent marker, biotin, or other detectable moiety known in the art that is detectable by analytical methods.

A detectable agent specific for a BRCA-1 regulator nucleic acid molecule can also be, for example, a PCR or RT-PCR primer, which can be used to selectively amplify all or a desired molecule, which can then be detected by methods known in the art. Furthermore, a detectable agent specific for a BRCA-1 regulator nucleic acid molecule can be a selective binding agent, such as a peptide, nucleic acid analog, or small organic molecule, identified, for example, by affinity screening of a library of compounds.

A detectable agent specific for a BRCA-1 regulator polypeptide can be, for example, an agent that selectively binds the polypeptide. For example, such a detectable agent is one that selectively binds with high affinity or avidity to the polypeptide without substantial cross-reactivity with other polypeptides that are not BRCA-l regulator polypeptides. The binding affinity of a detectable agent that selectively binds a polypeptide will generally be greater than about 10^{-5} M and more preferably greater than about 10^{-6} M for the polypeptide. High affinity interactions are preferred, and will generally be greater than about 10^{-8} M to 10^{-9} M.

A detectable agent specific for a BRCA-1 regulator polypeptide can be, for example, a polyclonal or monoclonal antibody specific for the polypeptide, or other selective binding agent identified, for example, by screening a library of compounds for binding to the regulator polypeptide. For certain applications, a detectable agent can be utilized that preferentially recognizes a particular conformational or post-translationally modified state of the polypeptide. The detectable agent can be labeled with a detectable moiety, if desired, or rendered detectable by specific binding to a detectable secondary binding agent.

The invention provides a substantially pure nucleic acid comprising a nucleotide sequence greater than about 57% identical to SEQ ID NO: 1, or a unique fragment thereof. Also provided is a substantially pure polypeptide comprising an amino acid sequence greater than about 41% identical to SEQ ID NO: 2, or functional fragment thereof.

The nucleic acid shown as SEQ ID NO: 1 has been found to encode a BRCA-1 regulator, which functions in reducing the expression of BRCA-1. As such, it is useful as a target for treating or reducing the severity of a cancer by regulating its proliferation through modulating BRCA-1 expression. Inhibition of the expression or activity of a BRCA-1 regulator results in a concomitant increase in BRCA-1 expression and therefore will result in decreased tumor cell proliferation.

SEQ ID NO: 1 corresponds to the expressed message of the human gene encoding BRCA-1 regulator termed BRCA-1 regulator 1 (BR1). SEQ ID NO: 1 is about 2967 nucleotides in length and has 5' and 3' non-coding regions of 395 and 988 nucleotides, respectively (Figure 9). The resultant coding region is 1584 nucleotides in length, coding for a polypeptide of 521 amino acids. SEQ ID NO: 1 has a nucleotide sequence of about 57% identical to the human EFG sequence.

Modifications of SEQ ID NO: 1, which do not substantially affect the activity of the encoded BRCA-1 regulator and which maintain nucleotide sequence identity greater than about 57% are included as nucleic acids of the invention. These nucleic acids having minor modifications can similarly be used for the development of therapeutic compounds which inhibit the expression or activity of BR1. Such modifications include, for example, changes in the nucleotide sequence which do not alter the encoded amino acid sequence as well as changes in the nucleotide sequence resulting in conservative amino acid substitutions or minor alterations which do not substantially affect the BRCA-1 regulating activity of BR1. Those skilled in the art will known or can determine what changes within greater than about 57% compared to SEQ ID NO: 1 can be made without substantially affecting the activity of BR1 as a regulator of BRCA-1 expression or activity.

Unique fragments of SEQ ID NO: 1 are also provided. The fragments are useful in a variety of procedures, including for example, as probes for determining the effectiveness of therapeutic agents which target expression of BR1 as a regulator of BRCA-1 expression or activity. Unique fragments also can be used to encode functional fragments of BR1 as therapeutic targets for anti-cancer compounds in the screening methods of the invention. The unique fragments of SEQ ID NO:1 are applicable in a variety of other methods and procedures known to those skilled in the art.

Unique fragments of SEQ ID NO: 1 correspond to fragments or portions of SEQ ID NO: 1 that are of sufficient length to distinguish the fragment as a BR1 encoding nucleic acid and that contain at least one nucleotide characteristic of SEQ ID NO: 1. Such a characteristic nucleotide, or nucleotides, within a specific fragment of SEQ ID NO: 1 distinguish that fragment from other related nucleotide sequences. For example, fragments of the non-coding region of SEQ ID NO: 1 are generally unique when compared to nucleotide sequences such as EFG, for example, because there is little evolutionary pressure to conserve non-coding domains. Nucleic acid sequences as small as between about 12-15 nucleotides are statistically unique sequences within the human genome. However, nucleic acids as small as between about 8-12 nucleotides can be unique. Therefore, non-coding region fragments of SEQ ID NO: 1 of about 8-9, preferably about 10-11, and more preferably about 12 or 15 nucleotides or more in length can be nucleotide sequences corresponding to unique fragments of SEQ ID NO: 1 of the invention.

Additionally, unique nucleotide sequences arise in the coding region of SEQ ID NO: 1 as well. Those skilled in the art will know or can determine which nucleotide positions are unique to either a non-coding region or a coding region fragment of SEQ ID NO: 1 given the teachings described herein or by alignment of SEQ ID NO: 1 with other sequences to be distinguished using methods well known to those skilled in the art. The nucleic acid sequence alignment of SEQ ID NO: 1 with the nucleic acid sequence of EFG shows that there are about 285 codon differences between the BR1 and EFG sequences indicating at least about 694 or more unique nucleotides in SEQ ID NO: 1 compared to the EFG encoding nucleotide sequence (Figure 2). Inclusion of any or all of these distinguishing nucleotide differences within a fragment of SEQ ID NO: 1 confers uniqueness onto the fragment.

A substantially pure nucleic acid molecule having a nucleotide sequence greater than about 57% identical to SEQ ID NO: 1, or a unique fragment thereof, will be of sufficient length and identity to SEQ ID NO: 1 to selectively hybridize to it under at least moderately stringent hybridization conditions. For example, it can be determined that a substantially pure nucleic acid molecule contains a nucleotide sequence greater than about 57% as SEQ ID NO: 1, or a unique fragment thereof, by

determining its ability to hybridize in a filter hybridization assay to a molecule having the sequence of SEQ ID NO: 1, but not to other unrelated nucleic acid molecules, under conditions equivalent to hybridization in 50% formamide, 5X Denhart's solution, 5X SSPE, 0.2% SDS at 42°C, followed by washing in 0.2X SSPE, 0.2% SDS, at 65°C. Suitable alternative buffers and hybridization conditions that provide for moderately stringent hybridization conditions in particular assay formats are known or can be determined by those skilled in the art (see, for example, Sambrook et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989).

The nucleic acid shown as SEQ ID NO: 1 encodes a polypeptide having the amino acid sequence shown in SEQ ID NO: 2. As with the nucleotide sequence described above, modifications of SEQ ID NO: 2 which do not substantially affect the activity of the BRCA-1 regulator and which maintain amino acid sequence identity greater than about 41% are included as polypeptides of the invention. These polypeptides having minor modifications can similarly be used for the development of therapeutic compounds which inhibit the BRCA-1-regulating activity of BR1. Such modifications include, for example, changes in the amino acid sequence which do not substantially alter the structure or function of a domain within the polypeptide as well as changes in the amino acid sequence which result in conservative substitutions or minor alterations which do not substantially affect the activity of BR1. Those skilled in the art will known or can determine what changes within greater than about 41% compared to SEQ ID NO: 2 can be made without substantially affecting the BRCA-1-regulating activity of BR1.

Functional fragments of SEQ ID NO: 2 are also provided. The BRCA-1 regulator BR1 can regulate BRCA-1 by inhibiting the translation of BRCA-1 mRNA. Therefore, it is within the scope of the invention that an example of a functional fragment of BR1 is a domain that inhibits translation of a mRNA into a polypeptide, for example the, translation of BRCA-1 mRNA into a polypeptide.

The invention also provides a substantially pure TST nucleic acid including a fragment having substantially the nucleotide sequence of SEQ ID NO: 1, wherein the TST has substantially the nucleotide sequence 5'–N_5–GUC–N_8–3' or 5'–N_5–GUA–N_8–3' (SEQ ID NOS: 3 and 4, respectively). The TST nucleic acid portion of the

fragment can have between about 8-12 nucleotides, and preferably about 9-10 nucleotides at positions N_5 and N_8, that are identical to a fragment of SEQ ID NO: 1. Therefore, depending on the length of the TST nucleic acid portion, a fragment of SEQ ID NO: 1 as described above can be about 11-15 nucleotides or greater in length.

Hairpin ribozymes cleave RNA substrates 5' to the G nucleotide in either of the sequences 5'–N_5–GUC–N_8–3' or 5'–N_5–GUA–N_8–3' (SEQ ID NOS: 3 and 4, respectively) Fragments of SEQ ID NO: 1 having a corresponding RNA form that is recognized by the target recognition site of a hairpin ribozyme therefore include nucleic acids having the sequence 5'–N_5–GUC–N_8–3' or 5'–N_5–GUA–N_8–3' (SEQ ID NOS: 3 and 4, respectively) where N_5 and N_8 are nucleotide sequences substantially the same as a sequence corresponding to SEQ ID NO: 1. Such fragments correspond to the complement sequence of a ribozyme target recognition site or RST, and are referred to herein as TST nucleic acids.

The TST nucleic acids can be of any desired length and can include additional sequences other than those corresponding to SEQ ID NO: 1 and other moieties so long as they have the structure 5'–N_5–GUC–N_8–3' or 5'–N_5–GUA–N_8–3' (SEQ ID NOS: 3 and 4, respectively) where N_5 and N_8 correspond substantially to a nucleotide sequence of SEQ ID NO: 1. Moreover, it is not necessary for all nucleotide residues within the N_5 and N_8 regions to be identical to the corresponding sequence within SEQ ID NO: 1. Instead, all that is necessary is for such TST nucleic acids to selectively hybridize to a complementary RST. Therefore, less than all 13 nucleotides at positions N_5 and N_8 can be identical to a nucleotide sequence or fragment of SEQ ID NO: 1. Generally, between about 8-12 or between about 9-10 nucleotides are sufficient for selective hybridization of a RST with a TST nucleic acid. Described further below in the Examples is a specific example of an RST present in a ribozyme that selectively hybridizes to SEQ ID NO: 1.

Similarly, the TST nucleic acids can be used to design ribozymes that selectively hybridize and cleave a RNA corresponding to SEQ ID NO: 1. A specific example of such a ribozyme is a hairpin ribozyme having a target recognition sequence complementary to a TST nucleic acid of SEQ ID NO: 1 and having the nucleotide sequence 5'–N_8–AGAA–N_4–3'. As with the TST nucleic acids described

above, it is not necessary that all the sequences with position N_8 and N_4 be identical in complement to SEQ ID NO: 1 so long as the target recognition sequence can selectively hybridize to the RNA form of TST nucleic acid and cleave it as a substrate. Those skilled in the art will know or can determine given the teachings and descriptions herein what RST sequences are sufficient for selective hybridization as well as for cleavage of a target RNA substrate.

Exemplary ribozymes selective for SEQ ID NO: 1 were constructed as described in the Examples. These ribozymes have the RST sequences: 5'–UUGAUGUGAGAAGCUU–3', 5'–ACUUUUCUAGAAGGAA–3', 5'–UAUUCCAUAGAAACUG–3', 5'–AGGACUGGAGAAAGCC–3', 5'–AACAACAUUAGAAUCAA–3' (SEQ ID NOS: 17-21).

Therefore, the invention provides a ribozyme having a target recognition sequence capable of selectively hybridizing to a RNA corresponding to SEQ ID NO: 1 and cleaving said RNA. The target recognition sequence of the ribozyme can consist of a RST complementary to a fragment of SEQ ID NO: 1 and having substantially the nucleotide sequence 5'–N_8–AGAA– N_4–3' (SEQ ID NO: 99). The target recognition sequence can further be between about 8-12 nucleotides, preferably about 9-10 nucleotides at positions N_8, and N_4 that are complementary to a fragment of SEQ ID NO: 1.

The invention further provides a RST having one of the following nucleotide sequences:

5'–CCGGAUGCAGAACAAU–3', 5'–AGUACAUUAGAAUACU–3' 5'–CUAGUGAGAGAAGGGA–3', 5'–UGAGAUCCAGAAAAGC–3' 5'–UGUUACUAGAAUGUU–3', and 5'–CCCUAUUUAGAAUUGU–3' (SEQ ID NOS: 5–10, respectively), or a complementary sequence thereof as described further below. The complementary sequence at positions 9-12 (5'–AGAA–3') can be substituted by the non-complementary ribozyme cleavage sequence 5'–NGUC–3'.

Libraries of ribozymes containing different RST sequences were expressed in cells, and cells expressing an increased level of BRCA-1 marker were selected. SEQ ID NOS: 5–10 represent RST sequences of ribozyme binding sites which, when the ribozymes were expressed, resulted in an increased level of BRCA-1 marker expression. BRCA-1 marker expression was determined by increased EGFP

expression, increased hygromycin resistance or decreased ability to grow in soft agar. These assays for BRCA-1 marker expression can be affected by both BRCA-1 promoter activity and BRCA-1 activity, since BRCA-1 itself can modulate BRCA-1 promoter activity. Therefore, the ribozymes containing RST sequences selective for a BRCA-1 regulator, such as the RSTs of SEQ ID NOS: 5–10, can increase BRCA-1 marker expression by selectively cleaving RNAs that regulate BRCA-1 promoter activity, BRCA-1 expression, BRCA-1 activity, or any combination thereof.

Each of the RST sequences SEQ ID NOS: 5–10 correspond to a nucleic acid encoding a BRCA-1 regulator As described previously with respect to TST nucleic acid fragments of the BRCA-1 regulator shown as SEQ ID NO: 1, the 5' terminal N_8 positions and the 3' terminal N_5 positions are separated by the intervening trinucleotide sequence 5'–GUC–3' in the BRCA-1 regulator RNA and correspond to the complement of the RST sequence or the TST sequence. Also as described previously, at least about 8-12 nucleotides within positions N_8 and N_5 of a TST are sufficient for selective binding between a ribozyme and its BRCA-1 regulator target RNA. Therefore, a BRCA-1 regulator nucleic acid molecule of the invention contains least about 8-12 nucleotides corresponding to a RST sequence set forth as SEQ ID NOS: 5–10, or its TST sequence complement corresponding to SEQ ID NOS: 11–16, and including the intervening trinucleotide 5'–GUC–3'.

Therefore, the invention also provides a TST nucleic acid having one of the following nucleotide sequences:
5'–AUUGNGUCGCAUCCGG–3', 5'–AGUANGUCAAUGUACU–3',
5'–UCCCNGUCCUCACUAG–3', 5'–GCUUNGUCGGAUCUCA–3',
5'–AACANGUCAGUAACA-3', and 5'–ACAANGUCAAAUAGGG-3'
(SEQ ID NOS: 11–16, respectively), or a complementary sequence thereof.

For simplicity of the description, the BRCA-1 regulator nucleic acids of the invention will be described with reference to its TST nucleic acid sequence and specifically with reference to a BRCA-1 nucleic acid containing at least about 8-12 regulator nucleotides corresponding to a TST nucleic acid sequence of SEQ ID NOS: 11–16. However, it is to be understood that reference to a BRCA-l regulator TST nucleotide sequence also specifically includes reference to the complementary sequence of the RST nucleic acid molecule. Therefore, it is also to be understood

that reference to a BRCA-1 regulator containing at least about 8-12 nucleotides corresponding to a TST sequence of SEQ ID NOS: 11–16 also includes reference to a BRCA-1 regulator nucleic acid containing at least about 11-15 nucleotides corresponding to a TST sequence set forth as SEQ ID NOS: 11–16, which includes the intervening trinucleotide 5'–GUC–3' between the 5' terminal N_8 positions and the 3' terminal N_5 positions and as described previously with respect to the BRCA-1 regulator shown as SEQ ID NO: 1 and its TST fragments. Therefore, the invention also provides substantially pure TST nucleic acids having the structure 5'–N_5–GUC–N_8–3' or 5'–N_5–GUA–N_8–3' and substantially the nucleotide sequences shown as SEQ ID NOS: 11–16.

A BRCA-1 regulator nucleic acid molecule containing at least about 8-12 nucleotides corresponding to a TST nucleic acid sequence set forth as SEQ ID NOS: 11–16, or about 11-15 nucleotides corresponding to a TST nucleic acid sequence set forth as SEQ ID NOS: 11–16, and including the intervening trinucleotide 5'–GUC–3' or a functional fragment thereof, does not have the exact endpoints of nucleotide sequences deposited and available in public databases. Such databases include, for example, the nonredundant GenBank database and NCBI dbest EST database.

A BRCA-1 regulator nucleic acid molecule of the invention containing at least about 8-12 nucleotides corresponding to a TST sequence set forth as SEQ ID NOS: 11–16, can be advantageously used, for example, as a therapeutic target for the treatment of cancer, as a diagnostic target for determining the presence or propensity for cancer, or to identify and isolate additional sequences corresponding to other regions of the BRCA-1 regulator nucleic acid molecules of the invention. When used for the latter purpose, the nucleic acid molecule can contain none, one, or many nucleotides at the 5' or 3' end, or both, of the TST sequences recited as SEQ ID NOS: 11–16. These additional nucleotides can correspond to the native sequence of the BRCA-1 regulator nucleic acid molecule, or can be nonnative sequences, or both. For example, non-native flanking sequences that correspond to a restriction endonuclease site or a tag, or which stabilize the nucleic acid containing at least about 8-12 nucleotides corresponding to a TST nucleic acid sequence set forth as SEQ-ID NOS: 11–16, in a hybridization assay, can be advantageous when the

nucleic acid molecule is used as a probe or primer to identify or isolate longer BRCA-1 regulator nucleic acid molecules.

Native BRCA-1 regulator nucleotide sequences flanking the at least about 8-12 nucleotides corresponding to a TST sequence set forth as SEQ ID NOS: 11–16, can be determined by methods known in the art, such as RT-PCR, 5' or 3' RACE, screening of cDNA or genomic libraries, and the like, using an oligonucleotide having at least about 8-12 nucleotides corresponding to the TST sequence of SEQ ID NOS: 11–16 as a primer or probe, and sequencing the resultant product. The appropriate source of template RNA or DNA for amplification, extension or hybridization screening can be determined by those skilled in the art.

A specific example of a substantially pure BRCA-1 regulator nucleic acid molecule containing at least about 8-12 nucleotides of a TST corresponding to SEQ ID NO: 13 and flanking coding sequence is the BRCA-1 regulator nucleic acid molecule having the nucleotide sequence set forth as SEQ ID NO: 1. The isolation of SEQ ID NO: 1, based on knowledge of the corresponding RST sequence of SEQ ID NO: 7, is described further below in the Examples. Therefore, such procedures can be used to identify and substantially purify longer nucleic acid molecules that contain at least about 8-12 nucleotides corresponding to a TST of SEQ ID NOS: 11–16. Such molecules and their functional fragments can be used to produce BRCA-1 regulator polypeptides and specific antibodies, for example, by methods known in the art and described herein, for use in the diagnostic and therapeutic methods described herein and known in the art.

In this regard, SEQ ID NOS: 5, 6, 8 and 9, have similarly been used to identify flanking nucleic acid sequences of the corresponding encoded BRCA-1 regulators. Specifically, SEQ ID NO: 5 has been used to identify a BRCA-1 regulator which was found to correspond to breast basic conserved protein 1 (BBC1, GenBank accession number X64707). BBC1, identified previously from a human breast carcinoma cDNA library, is downregulated in hormone-refractory prostate cancer. BBC1 contains a 25 amino acid region which has strong similarity to the plant basic peptide P14.

SEQ ID NO: 5 also has been used to identify a protein which corresponds to human CHL1 related protein CHLR2, a partial sequence of which is reported in

GenBank as accession number U33834. A portion of CHLR2 has been reported to act as a helicase. CHLR2 localizes to the nucleus and is apparently expressed at high levels only in proliferating human cell lines.

SEQ ID NO: 8 has been used to identify a further BRCA-1 regulator which was found to correspond to a polypeptide termed inhibitor dominant negative 4 (ID4, GenBank accession number NM_001546). ID4 is a transcriptional repressor of basic helix-loop-helix transcription factors which prevents transcriptional activation by dimerization to transcription factors, thereby blocking the ability of the transcription factor to bind DNA. It is within the scope of the invention that ID4 can bind to a transcription factor that increases BRCA-1 expression, including BRCA-1.

SEQ ID NO: 8 also has been used to identify protein corresponding to ALL-1 fusion partner AF6 (GenBank accession number AB011399), a translation breakpoint sequence found in acute myeloid leukemia. AF6 also is considered to be a Ras-binding protein and regulator of cell junction formation.

SEQ ID NOS: 6 and 9 have been used to identify BRCA-1 regulators corresponding to polypeptides termed herein as BR2 and BR3 (GenBank accession numbers AL045940 and AI273697, respectively), which have been reported only as EST sequences, and have not been characterized.

SEQ ID NO: 6 also has been used to identify an EST, termed BR4, having GenBank accession No. AI668913.

As described previously, a BRCA-1 regulator nucleic acid molecule, when functionally inactivated in a cell, results in increased BRCA-1 expression or activity, or both. Such increase results in the concomitant decrease in the ability of a cell to proliferate or undergo uncontrolled growth. Similar results can be observed by inactivation of the BRCA-1 regulator polypeptide by, for example, inhibiting its activity. The BRCA-1 regulator activity of a nucleic acid molecule containing at least about 8-12 nucleotides corresponding to a TST of SEQ ID NOS: 11–16 and additional native nucleic acid sequences can be further demonstrated using various methods known in the art and described herein. For example, nucleic acid sequences flanking the SEQ ID NOS: 11–16 sequences can be selectively targeted in a cell with ribozymes by the methods described herein. The effect on cell proliferation due to decreased BRCA-1 expression or activity or both can be determined by the assays

described below. If inactivation by ribozymal cleavage of a second sequence within the isolated nucleic acid molecule also results in increased BRCA-1 expression or activity, that nucleic acid molecule is a confirmed BRCA-1 regulator nucleic acid molecule.

For example, ribozymes targeted toward a second sequence within the nucleic acid molecules of BBC1, ID4 and BR1, have increased BRCA-1 or reporter gene expression and/or decreased cell proliferation in soft agar, further confirming these genes as encoding BRCA-1 regulators.

Therefore, knowledge of the nucleic acid sequence of a BRCA-1 regulator provides information that can be used to target further ribozymes against the BRCA-1 regulator, as demonstrated in the Examples. For example, a portion of one of the above-described BRCA-1 regulators which has a sequence 5'–N_5–GUC– N_8–3' or 5'–N_5–GUA– N_8–3' is considered to be a TST. Such a TST can be used as a template to design novel ribozymes having RST sequences 5'–N_8–AGAA– N_4–3' where N_8 and N_4 have nucleotide sequences substantially complementary to the corresponding residues in the 5'–N_5–GUC– N_8–3' or 5'–N_5–GUA– N_8–3' TST sequences. The RST sequence can have between 8-12 nucleotides, and preferably 9-10 nucleotides at positions N_8 and N_4 that are complementary to the corresponding 5'–N_5–GUC– N_8–3' or 5'–N_5–GUA– N_8–3' TST sequences.

As described below in the Examples, additional ribozymes directed to a variety of TST sequences on each of the previously described BRCA-1 regulators have been constructed. Specifically constructed ribozymes were:

ribozymes selective for BBC1 having the RST sequences:
5'–GGCUUCAAAGAAAUGC–3', 5'–UGGGACCCAGAACGGG–3', 5'–CCGGAUGGAGAACGAC–3', 5'–ACGUUCCGAGAAGGCA–3', 5'–CCCAGCAUAGAAGCCC–3' (SEQ ID NOS: 22–26);

ribozymes selective for CHLR2 having the RST sequences:
5'–CCGAGAGAAGAAAGCC–3', 5'–UGGUUGGAAGAACCGA–3', 5'–GAGGAUGCAGAACCAC–3', 5'–AAGAAACAAGAAACCC–3', 5'–UUGGCCAGAGAAGGGG–3' (SEQ ID NOS: 27–31);

ribozymes selective for ID4 having the RST sequences:
5'–CAGUGGGCAGAACUCA–3', 5'–CCAACAAUUAGAAGGAG–3',

5'–CACACCUGAGAAGCGC–3', 5'–CGCGGCUGAGAAGGUC–3' (SEQ ID NOS: 32–35);

ribozymes selective for AF6 having the RST sequences: 5'–GUACUAGAAGAACGAA–3', 5'–UGUGAUCCAGAAAAGG–3', 5'–GGUGGCCAAGAAGUGG–3' (SEQ ID NOS: 36–38);

ribozymes selective for BR2 having the RST sequences: 5'–AAAAAUUAAGAAGUCA–3', 5'–GCUGUCCUAGAAUCAA–3', 5'–UGUCAAAGAGAACACC–3', 5'–UGCAAUGAAGAAACUG–3', 5'–UUACAAUAAGAAACUU–3' (SEQ ID NOS: 39–43);

and ribozymes selective for BR3 having the RST sequences: 5'–CUAUUUAAAGAAAAUU–3', 5'–UAUUUCUUAGAAGUUC–3', 5'–AUUUCACUAGAAUCAC–3' (SEQ ID NOS: 44–46).

Similarly, other types of methods can be used to corroborate the activity of a BRCA-1 regulator nucleic acid containing at least about 8-12 nucleotides of a TST corresponding to SEQ ID NO: 11–16. For example, an antibody or other selective agent that binds a polypeptide encoded by the nucleic acid molecule can be introduced into the cell, and the effect of the antibody on BRCA-1 expression-or activity determined, for example, by measuring the ability of the cells to proliferate in soft agar. Similarly, an antisense nucleic acid that inhibits transcription or translation of the BRCA-1 regulator nucleic acid can be introduced into a cell, and the effect of the antisense nucleic acid on BRCA-1 expression or activity determined. Likewise, an altered form of a BRCA-1 regulator nucleic acid molecule, such as a dominant-negative mutant, can be expressed in a cell and its encoded polypeptide will compete with or inhibit an endogenous BRCA-1 regulator molecule, and thus increase BRCA-1 expression or activity. Those skilled in the art can determine other appropriate assays to demonstrate that a substantially pure nucleic acid molecule containing, at least about 8-12 nucleotides of any of SEQ ID NOS: 11–16 have BRCA-1 regulator activity.

The TST sequences set forth as SEQ ID NOS: 11–16, were identified from a random hairpin ribozyme library by assessing the ability of their corresponding RST to increase BRCA-1 expression (SEQ ID NOS: 5–10). Therefore, the invention provides ribozymes containing the RST sequences set forth as SEQ ID NOS: 5–10 as

the ribozyme target recognition sequence. The hairpin ribozymes of the invention selectively bind to BRCA-1 regulator mRNA molecules complementary, in part, to these RST sequences.

A substantially pure hairpin ribozyme of the invention can be catalytic, so as to bind and cleave a BRCA-1 regulator nucleic acid messenger RNA. A catalytic hairpin ribozyme of the invention can therefore be used to selectively regulate the activity of a BRCA-1 regulator nucleic acid molecule of the invention. A substantially pure hairpin ribozyme of the invention can also be catalytically disabled, for example, by replacement of the Loop 2 AAA sequence indicated in Figure 1 with a UGC sequence, so as to bind, but not cleave, a BRCA-1 regulator nucleic acid molecule of the invention. A non-catalytic hairpin ribozyme can be used, for example, as a control for the inhibition activity of non-disabled ribozymes.

Therefore, the invention also provides a ribozyme containing a target recognition sequence having any one of the following nucleotide sequences: 5'–CCGGAUGCAGAACAAU–3', 5'–AGUACAUUAGAAUACU–3' 5'–CUAGUGAGAGAAGGGA–3', 5'–UGAGAUCCAGAAAAGC–3' 5'–UGUUACUAGAAUGUU–3', and 5'–CCCUAUUUAGAAUUGU–3' (SEQ ID NOS: 5–10, respectively).

Also provided herein is a method of identifying ribozyme cleavage targets from a nucleic acid sample by using the sequence specificity of the ribozyme to isolate a ribozyme cleavage product nucleic acid. This method can be used for cloning a flanking sequence of any target nucleic acid for which there is a ribozyme cleavage site. A nucleic acid sample is subjected to cleavage by a sequence-specific test ribozyme. The sequence specific recognition and cleavage by a ribozyme results in at least one cleavage product nucleic acid having a specific sequence at either the 5' or 3' end, which corresponds to the sequence recognized by the ribozyme. When the other end of the cleavage product nucleic acid contains a general sequence, for example, a 3' poly A tail, two primers can be constructed, a first complementary to the sequence recognized by the ribozyme, and a second complementary to the general sequence. These two primers can then be used to amplify the cleavage product nucleic acid, which then can be identified using known sequencing methods.

Thus from a nucleic acid sample, those nucleic acids that are cleaved by a ribozyme can be specifically amplified and subsequently identified.

Target nucleic acids can be either RNA or DNA, and can be cleaved by any ribozyme, including a hairpin ribozyme, a hammerhead ribozyme and a Tetrahymena group I ribozyme. Prior to sequencing, amplification products can be separated by gel electrophoresis and can further be compared to a nucleic acid sample subjected to cleavage by a control ribozyme having sequence specificity different than the ribozyme used in the nucleic acid sample being tested. The amplification products that are specific for the test ribozyme cleaved sample can be recovered and sequenced.

One embodiment of this method uses four primers. A first primer hybridizes to a 3' end first general sequence of the cleavage product nucleic acid, and strand synthesis is carried out. The enzyme used for such a synthesis places a second general sequence at the 3' end of the newly synthesized strand, which also is adjacent the specific flanking sequence of ribozyme cleavage. Such an exemplary enzyme is MMLV reverse transcriptase, which places 3-5 deoxycytidine nucleotides at the 3' end of the nascent DNA strand. Next, a second primer which contains a sequence complementary to the second general sequence is added, and the nascent nucleic acid is further extended. The template strand is then removed, for example, by RNAse digestion if the template strand is RNA. The remaining nucleic acid is then amplified using a third primer complementary to the sequence of the first primer and a fourth primer complementary to the sequence of the second primer and containing the flanking sequence of ribozyme cleavage. The amplified product is then identified using sequencing methods known in the art.

The nucleic acid molecules of the invention, including BRCA-1 regulator nucleic acid molecules and fragments, and hairpin ribozyme nucleic acid molecules, can be produced or isolated by methods known in the art. The method chosen will depend, for example, on the type of nucleic acid molecule one intends to isolate. Those skilled in the art, based on knowledge of the nucleotide sequences described herein, can readily isolate BRCA-1 regulator nucleic acid molecules as genomic DNA, or desired introns, exons or regulatory sequences therefrom; as full-length cDNA or desired fragments therefrom; or as full-length mRNA or desired fragments

therefrom, by methods known in the art. Likewise, those skilled in the art can produce or isolate hairpin ribozymes selective for these sequences.

A useful method of isolating a BRCA-1 regulator nucleic acid molecule of the invention involves amplification of the nucleic acid molecule using the polymerase chain reaction (PCR), and purification of the resulting product by gel electrophoresis. For example, either PCR or reverse-transcription PCR (RT-PCR) can be used to produce a BRCA-1 regulator nucleic acid molecule having any desired nucleotide boundaries. Desired modifications to the nucleic acid sequence can also be introduced by choosing an appropriate primer with one or more additions, deletions or substitutions. Such nucleic acid molecules can be amplified exponentially starting from as little as a single gene or mRNA copy, from any cell, tissue or species of interest. An example of the isolation of a BRCA-1 regulator nucleic acid molecules using PCR are presented below in the Examples.

Another method of producing or isolating a BRCA-1 regulator nucleic acid molecule of the invention is by screening a library, such as a genomic library, cDNA library or expression library, with a detectable agent. Such libraries are commercially available or can be produced from any desired tissue, cell, or species of interest using methods known in the art. For example, a cDNA or genomic library can be screened by hybridization with a detectably labeled nucleic acid molecule having a nucleotide sequence disclosed herein. Additionally, an expression library can be screened with an antibody raised against a polypeptide corresponding to the coding sequence of a BRCA-1-regulator nucleic acid disclosed herein. The library clones containing BRCA-1 regulator nucleic acid molecules of the invention can be purified away from other clones by methods known in the art.

Furthermore, nucleic acid molecules of the invention can be produced by synthetic means. For example, a single strand of a nucleic acid molecule can be chemically synthesized in one piece, or in several pieces, by automated synthesis methods known in the art. The complementary strand can likewise be synthesized in one or more pieces, and a double-stranded molecule made by annealing the complementary strands. Direct synthesis is particularly advantageous for producing relatively short molecules, such as RST or hairpin ribozyme nucleic acid molecules, as well as hybridization probes and primers. Alternatively, nucleic acid molecules

of the invention can be produced using in vitro enzymatic synthesis using RNA polymerase. For example, a template DNA can be combined with RNA polymerase and ribonucleotide triphosphates. The RNA polymerase synthesizes an RNA nucleic acid complementary to the sequence of the DNA template.

If it is desired to subclone, amplify or express a substantially pure nucleic acid molecule of the invention, the isolated nucleic acid molecule can be inserted into a commercially available cloning or expression vector using methods known in the art. Appropriate regulatory elements can be chosen, if desired, to provide for constitutive, inducible or cell type-specific expression in a host cell of choice, such as a bacterial, yeast, amphibian, insect or mammalian cell, including human cells. Those skilled in the art can determine an appropriate host and vector system for cloning a nucleic acid molecule of the invention or for expressing and purifying its encoded polypeptide.

Methods for introducing a cloning or expression vector into a host cell are well known in the art and include, for example, various methods of transfection such as the calcium phosphate, DEAE-dextran and lipofection methods, viral transduction, electroporation and microinjection. Host cells expressing BRCA-1 regulator nucleic acid molecules can be used, for example, as a source to isolate recombinantly expressed BRCA-1 regulator polypeptides, to identify and isolate molecules that regulate or interact with BRCA-1 regulator nucleic acids and polypeptides, or to screen for compounds that enhance or inhibit the activity of a BRCA-1 regulator molecule of the invention, as described further below.

The methods of isolating, cloning and expressing nucleic acid molecules of the invention described herein are routine in the art and are described in detail, for example, in Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York (1992) and in Ansubel et al., Current Protocols in Molecular Biology, John Wiley and Sons, Baltimore, MD (1989) .

BRCA-1 regulator polypeptides and functional fragments of the invention can be isolated or prepared by methods known in the art, including biochemical, recombinant and synthetic methods. For example, a BRCA-1 regulator polypeptide can be purified by routine biochemical methods from a cell or tissue source that expresses abundant amounts of the corresponding transcript or polypeptide.

Biochemical purification can include, for example, steps such as solubilization of the appropriate tissue or cells, isolation of desired subcellular fractions, size or affinity chromatography, electrophoresis, and immunoaffinity procedures. The methods and conditions for biochemical purification of a polypeptide of the invention can be chosen by those skilled in the art, and purification monitored, for example, by an ELISA assay or a functional assay.

A fragment having any desired boundaries and modifications to a BRCA-1 regulator amino acid sequences can also be produced by recombinant methods. Recombinant methods involve expressing a nucleic acid molecule encoding the desired polypeptide or fragment in a host cell or cell extract, and isolating the recombinant polypeptide or fragment, such as by routine biochemical purification methods described above. To facilitate identification and purification of the recombinant polypeptide, it can be desirable to insert or add, in-frame with the coding sequence, nucleic acid sequences that encode epitope tags, polyhistidine tags, glutathione-S-transferase (GST) domains, and similar affinity binding sequences, or sequences that direct expression of the polypeptide in the periplasm or direct secretion. Methods for producing and expressing recombinant polypeptides *in vitro* and in prokaryotic and eukaryotic host cells are well known in the art.

Functional fragments of a BRCA-1 regulator polypeptide can also be produced, for example, by enzymatic or chemical cleavage of the full-length polypeptide. Methods for enzymatic and chemical cleavage and for purification of the resultant peptide fragments are well known in the art (see, for example, Deutscher, Methods in Enzymology, Vol. 182, "Guide to Protein Purification," San Diego: Academic Press, Inc. (1990)), which is incorporated herein by reference.

Furthermore, functional fragments of a BRCA-l regulator polypeptide can be produced by chemical synthesis. If desired, such as to optimize their functional activity, stability or bioavailability, such molecules can be modified to include D-stereoisomers, non-naturally occurring amino acids, and amino acid analogs and mimetics. Examples of modified amino acids and their uses are presented in Sawyer, Peptide Based Drug Design, ACS, Washington (1995) and Gross and Meienhofer, The Peptides: Analysis, Synthesis, Biology, Academic Press, Inc., New York (1983), both of which are incorporated herein by reference.

A functional activity of a BRCA-1 regulator polypeptide or fragment of the invention can be its ability to alter BRCA-1 expression or activity, for example increasing unrestricted cell growth, when expressed or introduced in a cell. To determine whether a given polypeptide or fragment has the ability to alter BRCA-1 expression or activity, a polypeptide or fragment can be expressed in the cell by recombinant methods known in the art and the effect of the BRCA-1 regulator can be determined *in vitro*. Alternatively, expression of the BRCA-1 regulator can be inhibited *in vivo*, including cell culture or animal models and BRCA-l expression or activity can be determined. Similarly, expression of the BRCA-1 regulator can be inhibited *in vivo*, including cell culture or animal models and the expression of a BRCA-1 promoter-linked reporter marker determined. An increase in cell proliferation in soft agar or decrease in the expression of a BRCA-1 promoter-linked reporter marker indicates that the polypeptide or fragment is a BRCA-1 regulator of the invention.

The invention provides a method of identifying a compound that modulates the activity of a BRCA-1 regulator comprising contacting a BRCA-1 regulator with a target molecule responsive to the activity of the BRCA-1 regulator and the test compound. In this method, an increase or decrease in the activity of the BRCA-1 regulator for the target molecule compared to the absence of the test compound indicates that the compound modulates the activity of the BRCA-1 regulator. As used herein, "a target molecule responsive to the activity of the BRCA-1 regulator" means that the BRCA-1 regulator either binds or chemically modifies the a target molecule that exists in a cell. For example, if the BRCA-1 regulator has DNA binding activity for use in, for example, transcriptional gene regulation, the target molecule responsive to this activity is a nucleic acid comprising a sequence of nucleotides recognized and bound by the particular BRCA-1 regulator. If the BRCA-1 regulator has protein kinase activity, due for example to having a serine/ threonine kinase domain, the target molecule responsive to this activity is a protein having an appropriate serine or threonine kinase recognition site. Likewise, for activity as a protease, the target molecule responsive is a protein cleaved by the BRCA-1 regulator.

When the BRCA-1 regulator activity to be measured for drug screening is DNA binding, such binding can be determined by assaying the expression of a reporter gene that is operatively linked to the nucleic acid element. In this case, an increase in the amount of expression or activity of the reporter gene in the presence of a test compound compared to the absence of the test compound indicates that the compound has BRCA-1 regulator DNA binding inhibitory activity. The magnitude of the increase in expression activity will correlate with the BRCA-1 regulator inhibitory activity of the test compound. Exemplary reporter genes include BRCA-1, EGFP and hygromycin resistance gene.

As used herein, the term "nucleic acid element" when used in reference to regulation of BRCA-1 expression refers to a nucleic acid region that modulates BRCA-1 expression. Exemplary nucleic acid elements are the BRCA-1 5' promoter and regulatory region or other transcriptional regulation regions, and translational regulatory regions of the transcribed mRNA. Generally, the nucleic acid element will be the 5' promoter and regulatory region.

Similarly, compounds that increase or enhance the activity of BRCA-1 regulator also can be identified. A test compound added to a sample containing a BRCA-1 regulator and a nucleic acid element modulated by a BRCA-1 regulator which decreases BRCA-1 activity or the amount or rate of expression of BRCA-1 or a reporter gene operatively linked to the nucleic acid element compared to the absence of the test compound indicates that the compound increases the activity of the BRCA-1 regulator. Therefore, the invention provides a method of identifying compounds that modulate the activity of a BRCA-1 regulator.

A reaction system for identifying a compound that inhibits or increases BRCA-l regulator activity can be prepared using essentially any sample, material or components thereof that contains a BRCA-l regulator. A BRCA-l regulator containing sample used for such methods can be, for example, in vitro transcription or translation systems using, for example, nucleic acid derived from the BRCA-1 gene-of a normal or tumor cell or a hybrid construct linking the nucleic acid element modulated by a BRCA-1 regulator to a reporter gene. Alternatively, nucleic acids and proteins obtained from normal cells can also be used since BRCA-1 regulators can also act in normal cells. The BRCA-l regulator-containing sample can

additionally be derived from cell extracts, cell fractions or, for example, *in vivo* systems such as cell culture or animal models which contain a nucleic acid element modulated by a BRCA-l regulator. The expression levels or activity of BRCA-l or the reporter gene can be measured in the reaction system to determine the modulatory effect of the test compound on the BRCA-1 regulator. Such measurements can be determined using methods described herein as well as methods well known to those skilled in the art.

Briefly, the BRCA-1 regulator source is combined with a nucleic acid element or protein modulated by a BRCA-1 regulator as described above and incubated in the presence or absence of a test compound. The expression levels or activity of BRCA-1 or the reporter gene in the presence of the test compound is compared with that in the absence of the test compound. Those test compounds which provide an increase in expression levels or activity of BRCA-1 or the reporter gene of at least about 50% are considered to be BRCA-1 regulator inhibitors, or antagonists, and further, potential therapeutic compounds for the treatment of neoplastic diseases such as cancer. Similarly, those compounds which decrease expression levels or activity of BRCA-1 or the reporter gene by about two-fold or more are considered to be compounds which increase the activity of a BRCA-1 regulator, or BRCA-1 regulator agonists. Such agonists can be used as therapeutics, for example, to promote cell growth or cell survival in transplanted or explanted cells which are subsequently transplanted. Compounds identified to modulate BRCA-1 regulator activity can, if desired, be subjected to further *in vitro* or *in vivo* studies to corroborate that they affect the activity of a BRCA-1 regulator toward the BRCA-1 expression or activity.

Suitable test compounds for the above-described assays can be any substance, molecule, compound, mixture of molecules or compounds, or any other composition which is suspected of being capable of inhibiting BRCA-1 regulator activity *in vivo* or *in vitro*, for example, compounds with cell proliferation-inhibiting activity. The test compounds can be macromolecules, such as biological polymers, including proteins, polysaccharides and nucleic acids. Sources of test compounds which can be screened for BRCA-1 regulator inhibitory activity include, for example, libraries of small organic molecules, peptides, polypeptides, DNA, and RNA. Additionally, test

compounds can be pre-selected based on a variety of criteria. For example, suitable test compounds can be selected as having known inhibition or enhancement activity with respect to cell proliferation. Alternatively, the test compounds can be selected randomly and tested by the screening methods of the present invention. Test compounds can be administered to the reaction system at a single concentration or, alternatively, at a range of concentrations to determine, for example, the optimal modulatory activity toward the BRCA-l regulator.

The activity of a BRCA-1 regulator for which drug screening is desired can be a protein kinase activity. For example, BRCA-1 regulators that have a serine/threonine kinase domain may be used for drug screening where the activity which is modulated is a protein kinase activity. Protein kinase assays are well known to those skilled in the art (see, *e.g.*, U.S. Pat. Nos. 5,538,858 and 5,757,787; Anal. Biochem, 209:348-353, (1993)).

The activity of a BRCA-1 regulator for which drug screening is desired also can be GTP binding activity. For example, BRCA-1 regulators that have a GTP binding site may be used for drug screening where the activity which is modulated is GTP binding. BRCA-1 regulators that have GTP-binding activity may regulate cell growth such as through regulating BRCA-1 expression, or may have affects on cell cycle control, protein secretion, and intracellular vesicle interaction. GTP binding assays are well known to those skilled in the art (see, *e.g.*, U.S. Pat. Nos. 5,840,969 to Hillman et al.).

The activity of a BRCA-1 regulator for which drug screening is desired also can be hormone binding activity. For example, BRCA-1 regulators that have a hormone binding site may be used for drug screening where the activity which is modulated is hormone binding. Hormones that bind to a BRCA-1 regulator may be steroid hormones such as estrogen or a protein based hormone. Receptor hormone binding assays including receptor estrogen binding assays are well known to those skilled in the art (see, *e.g.*, U.S. Pat. Nos. 6,204,067 to Simon et al.).

The invention provides a method of identifying a ribozyme reactive with a BRCA-1 regulator. The method comprises: (a) introducing a randomized ribozyme

library into a population of cells expressing a selection marker gene operatively linked to a nucleic acid element modulated by a BRCA-1 regulator; (b) subjecting the population of cells to selection, and (c) recovering one or more ribozymes from viable cells following selection.

Also provided is a method of identifying a BRCA-1 regulator. The method comprises: (a) introducing a randomized ribozyme library into a population of cells expressing a selection marker operatively linked to a nucleic acid element modulated by a BRCA-1-regulator; (b) subjecting the population of cells to selection; (c) recovering one or more ribozymes from viable cells following selection; and (d) sequencing the target recognition sequence of the recovered ribozyme to identify the nucleic acid encoding the BRCA-1 regulator.

By reference to cancer, or to breast or ovarian cancer in particular, as an exemplary neoplastic disease amenable to the methods of identifying a ribozyme or a BRCA-l regulator of the invention, one skilled in the art will readily know, in light of the teachings and description herein that such methods are applicable to essentially all neoplastic diseases which employ tumor suppressor regulators for continued propagation. Therefore, the methods of identifying a BRCA-1 regulator, or ribozyme selective to a BRCA-1 regulator, as well as methods of treating a neoplastic disease once such regulators have been identified are applicable to methods of identifying any tumor suppressor regulator and to methods of treating a neoplastic disease.

A method of identifying tumor suppressor regulators can be carried out using the methods described herein. Briefly, a nucleic acid element of a tumor suppressor gene that regulates the expression of the gene, for example, the 5' promoter and regulatory region for a specific tumor suppressor, can be operatively linked to a reporter gene. A sample containing a tumor suppressor regulator and a nucleic acid element modulated by the tumor suppressor regulator linked to a reporter gene can be contacted with a test compound and the expression or activity of the reporter gene can be measured. A test compound found to increase reporter gene expression or activity indicates that the test compound inhibits tumor suppressor regulator activity. A test compound found to decrease reporter gene expression or activity indicates that the test compound increases or enhances tumor suppressor regulator activity.

For the successful application of such methods, it is sufficient to have identified a nucleic acid, derived from a normal or tumor cell, which is modulated by a tumor suppress or regulator, for example, the BRCA-1 promoter or the p53 promoter. Once identified, the nucleic acid derived from the cell can be operatively linked to a reporter gene such as a selection marker gene, and subjected to selection using the methods of the invention.

Moreover, it is not necessary for the nucleic acid derived from the normal or tumor cell to be unique to the regulation of expression or activity of the specific tumor suppressor in that cell. Instead, the normal or tumor cell-derived tumor suppressor regulator nucleic acid can overlap or be redundant with other regulators of tumor suppressor expression or activity. The expression or activity level of the tumor suppressor can rely on a balance between levels of one or more tumor suppressor regulators found in normal cells compared to levels of one or more tumor suppressor regulators in tumor cells. Therefore, decreasing the level or activity of tumor suppressor regulators acting on common components or structures can shift the balance toward utilization of tumor suppressor regulators for normal cellular functions reducing the amount of unrestricted cell growth. An example of a nucleic acid modulated by a tumor suppressor regulator includes a nucleic acid that is expressed in normal cells, but whose expression is increased in tumor cells. Another example of a nucleic acid modulated by a tumor suppressor regulator includes a nucleic acid that is not expressed in normal cells, but is expressed in tumor cells.

The method of identifying a ribozyme reactive with a BRCA-1 regulator involves the construction of a population of cells expressing a selection marker gene which is under the control, or operatively linked to a nucleic acid modulated by a BRCA-1 regulator. Specific examples of such a cell population and its use are described further below in the Examples.

Briefly, a nucleic acid element modulated by a BRCA-1 regulator can be essentially any nucleic acid sequence that influences BRCA-1 expression or activity. Specific examples of such an element includes the BRCA-1 promoter and the 5' regulator region. Methods using BRCA-1 regulators identified for the BRCA-1 promoter and therapeutic compounds directed thereto are applicable to all neoplastic

diseases related to decreased BRCA-1 expression or activity. Other elements can include, for example, transcription enhancers and BRCA-1 mRNA.

Cell populations containing a nucleic acid element modulated by a BRCA-1 regulator are operatively linked to a selection marker gene. Operative linkage will depend on the type of element employed and is intended to refer to placing the nucleic acid element in an appropriate context and location in the reporter construct as it would be found in its native genome. In the specific example of BRCA-1 promoter, operative linkage places the element 5' to the transcription start site of the marker gene. Operative linkage of a promoter element-will be sufficiently upstream of the translation start codon to include sufficient 5' untranslated region sequence to effect translation in, for example, a CAP-dependent manner. The reporter constructs can be introduced into cell population using well known methods in the art and as described previously.

A selection marker can be a gene product that is, or can be made to influence cell viability or cell growth, or can be used as a measurable indicator. Specific examples include enhanced green fluorescence protein (EGFP), hygromycin resistance gene, the tumor suppressor BRCA-1, Herpes Simplex Virus thymidine kinase (HSV-tk), cytosine deaminase (CD) and diphtheria toxin (DT). For example, the expression of a negative selection marker in cells is either toxic alone, or toxic in the presence of a negative selection compound which is metabolized by the marker gene product into a cytotoxic or cytostatic substance. In contrast, expression of a positive selection marker in cells protects cells from arrested growth or cell death. Alternatively, the gene product can be a measurable indicator that can be used in selecting cells having a certain measured level of expression of this gene product.

Expression of an easily detectable gene product such as EGFP can be used as a measurable indicator of gene expression. EGFP allows a specific, user-defined expression level to be selected using methods such as fluorescence activated cell sorting (FACS). For example, cells representing the 10% of the population expressing the highest levels of EGFP can be recovered using FACS methods. The selected subpopulation can be grown and subjected to a second round of FACS-based selection of high level EGFP expression. This iterative FACS-based selection

results in enrichment of cells expressing high levels of EGFP and thus represents a FACS-based positive selection method.

Positive selection can also be carried out, for example, using hygromycin. Expression of the hygromycin resistance gene permits cells to survive in the presence of hygromycin. Hygromycin is an aminoglycoside antibiotic that kills bacteria, fungi and higher eukaryotic cells by inhibiting protein synthesis. Expression of hygromycin resistance gene allows cells to survive in the presence of hygromycin. Concentrations of hygromycin used for selection are between about 50 μg/ml and 1000 μg/ml, preferably about 250 μg/ml.

Phenotype selection can also be carried out. Expression of the BRCA-1 gene suppresses cell growth, and particularly anchorage-independent growth, resulting in, for example, a decreased ability to grow in soft agar. Growth of cells in soft agar can therefore be used to determine the propensity of the cells for anchorage-independent growth, thus indicating the level of BRCA-1 gene expressed.

Once selection proceeds, the selected cells are those which express a ribozyme that is reactive with a BRCA-1 regulator that inhibits or decreases BRCA-1 promoter-dependent expression of the reporter gene. The cells are isolated and the ribozymes are recovered using, for example, PCR or other well known methods in the art. The RST of the ribozyme is a sequence tag corresponding to a BRCA-1 regulator. Sequencing of this tag identifies the nucleic acid encoding the BRCA-1 regulator. Specific examples of RSTs corresponding to BRCA-1 regulators of the invention are set forth as SEQ ID NOS: 5–10 and their corresponding TSTs are set forth as SEQ ID NOS: 11–16. Five RST sequences have been determined. Specifically, SEQ ID NO: 7 corresponds to the RST for BR1, the full length nucleotide and amino acid sequences of which is shown as SEQ ID NOS: 1 and 2. SEQ ID NO: 5 corresponds to the RST for BBC1 (GenBank accession number X64707) and also for CHLR2 (GenBank accession number U33834) and SEQ ID NO: 8 corresponds to the RST for ID4 (GenBank accession number NM_001546) and also for AF6 (GenBank accession number AB011399). SEQ ID NO: 6 corresponds to the RST for BR2 (GenBank accession number AL045940), and SEQ ID NO: 9 corresponds to the RST for BR3 (GenBank accession number AI276397).

The invention also provides a method of treating breast cancer or ovarian cancer. The method consists of introducing a ribozyme selectively reactive with a RNA encoding BBC1, BR1, ID4, BR2, or BR3, into a cancerous breast or ovarian cell, preferably the ribozyme is selectively reactive with an RNA encoding BBC1, ID4 or BR1. Also provided is a method of treating breast or ovarian cancer by introducing a ribozyme selectively reactive with a RNA encoding a BRCA-1 regulator corresponding to a RST selected from the group consisting of SEQ ID NOS: 5–10. By substituting the ribozymes of the invention selectively reactive with a BRCA-1 regulator RNA with an antisense nucleic acid corresponding to a RST sequence selected from the group consisting of SEQ ID NOS: 5–10, methods of treating breast cancer or ovarian cancer are also provided. The antisense nucleic acids hybridize to the BRCA-1 regulator nucleic acid similar to catalytic ribozymes and inhibit transcription processing or translation of the RNA without subsequent cleavage. Such methods will be described below with reference to a ribozyme of the invention, but those skilled in the art will know that antisense nucleic acids can similarly be substituted for the ribozymes to prevent or reduce the severity of breast cancer or ovarian cancer.

A ribozyme encoding any of the RST sequences set forth as SEQ ID NOS: 5–10, or a combination thereof can be delivered in a wide variety of ways to tumor cells or cells susceptible to uncontrolled proliferation to interrupt or prevent neoplastic growth. The ribozyme can be administered as RNA or expressed from an expression vector. The ribozyme can be administered *ex vivo* by, for example, administering to cells that have been removed from an infected individual, and then returned to the individual, or the ribozyme can be administered *in vivo*. Delivery can be performed using any appropriate delivery vehicle known to those skilled in the art including, for example, a liposome, a controlled release vehicle, electroporation or covalently attached moieties, and other pharmacologically acceptable methods of delivery. A carrier can provide specificity for tissue or organ accumulation, such as breast or ovary accumulation, of the ribozyme at the tissue or organ which is the primary site of the neoplastic-growth. The ribozyme delivery vehicle can be designed to serve as a slow release reservoir or to deliver its contents directly to the target cell. Examples of ribozyme delivery vehicles include liposomes, hydrogels, cyclodextrins,

biodegradable nanocapsules, and bioadhesive microspheres. Liposomes can readily be targeted to the liver for delivery of RNA to infected tumor cells.

Routes of ribozyme administration include intramuscular, aerosol, intravenous, parenteral, intraperitoneal. Generally however, the route of administration will be through a blood vessel which represents a direct route-to the organ or tissue to be treated. The dosage of ribozyme will also depend on a variety of factors, such as the form of the ribozyme, the route of administration, the severity of the cancer or stage of disease, the general condition of the patient being treated, and thus can vary widely. Generally the dosage of ribozyme will be between about 10μg - 200mg/kg of body weight per day and result in therapeutic or prophylactic levels within the targeted cells sufficient to inhibit or eradicate the tumor cells. The duration of treatment can extend throughout the course of the neoplastic disease or disease symptoms, usually at least about 7-30 days, with longer duration being necessary for severe diseases. The number and timing of doses can also vary depending on, for example, the extent of disease.

A viral vector containing a ribozyme corresponding to a BRCA-1 regulator RST of the invention can be prepared in any of a wide variety of ways known to those skilled in the art. Representative retroviral vectors which can be used in the methods of the invention are described, for example, in U.S. Patent Nos. 4,861,719, 5,124,263 and 5,219,740. Other vectors can also be employed, particularly for the *ex vivo* methods, such as DNA vectors, pseudotype retroviral vectors, adenovirus, and adeno-associated virus vectors.

The viral vector contemplated for use in the methods of the present invention comprises infectious, but replication-defective, viral particles, which contain at least one DNA sequence encoding a ribozyme selectively reactive with a BRCA-1 regulator, is administered in an amount effective to inhibit or prevent cancer formation or progression in a subject. The vector particles can be administered in an amount-from 1 plaque forming unit to about 10^{14} plaque forming units, more preferably from about $1X10^{6}$ plaque forming units to about $1X10^{13}$ plaque forming units. A sufficient number of vector particles containing a ribozyme selectively reactive with a BRCA-1 regulator of the invention is administered to the cancerous tissue or organ to infect up to at least about 50% of the cancerous cells, usually about

80%, preferably about 90%, or more of the cancerous cells in the individual. Subsequent administrations can be performed, as needed, to effectively treat or reduce the severity of the cancer.

Exemplary ribozymes of the invention include, for example, those having RST sequences set forth as SEQ ID NOS: 5–10. One of these RST sequences, SEQ ID NO: 5, corresponds to the proteins BBC1 and CHRL2. Another RST sequence, SEQ ID NO: 8, corresponds to the proteins ID4 and AF6. Three of the RST sequences correspond to sequences without previously known roles. Briefly, SEQ ID NO: 7 corresponds to BR1 (SEQ ID NO: 1), SEQ ID NO: 6 corresponds to BR2 and SEQ ID NO: 9 corresponds to BR3.

Methods of treating cancer are also provided by this invention. More specifically, within one aspect of the present invention cancerous conditions may be treated by administering to a warm-blooded animal (e.g., a human) a therapeutically effective amount of ribozyme.

The present invention provides for treatment of cancer by contacting desired cells with an effective amount of ribozyme of this invention. A suitable "therapeutically effective amount" will depend on the nature and extent of diseased tissue being treated. Such "therapeutically effective amounts" can be readily determined by those of skill in the art using well known methodology, and suitable animal models such as a rat, rabbit, or porcine model, or, based upon clinical trials. As utilized herein, a patient is deemed to be "treated" if the cancerous condition is reversed or inhibited in a patient in a quantifiable manner. A therapeutically effective amount or regimen of treatment should result in: (1) decrease in the frequency, severity, or, duration of clinical symptoms (e.g., pain); (2) increase of time in the period of remission; (3) a change in pathological symptoms (e.g. reduction in tumor volume or tumor markers); or (4) any combination thereof. When exogenously delivering the ribozyme, the RNA molecule can be embedded within a stable RNA molecule or in another form of protective environment, such as a liposome. Alternatively, the RNA can be embedded within Rnase-resistant DNA counterparts. Cellular uptake of the exogenous ribozyme can be enhanced by attaching chemical groups to the DNA ends, such as cholesteryl moieties (Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989).

Hammerhead ribozymes suitable for use within the present invention preferably recognize the sequence NUH, wherein N is any of G, U, C, or A and H is C, U, or A. The hairpin recognition sites GUC are a subset of the hammerhead recognition sites NUH. Therefore, all hairpin sites are by definition also hammerhead sites, although the converse is not true. Chimeric hammerhead ribozymes (i.e., RNA/DNA hybrids) are designed to have an appropriate NUH sequence for ribozyme cleavage. Ribozymes are chemically synthesized with the general structure shown below as Scheme 1. The binding arms bases and stem loop comprise DNA, and the catalytic domain comprises RNA and/or 2' O methyl RNA bases. The stem loop can be replaced by a propanediol linker. DNA bases are shown in upper case, RNA bases in lower case, 2' O methyl RNA is shown as lower case underlined, and propanediol linker is shown as pr pr pr pr.

Scheme 1

Sequence ID No: 47 Length: 38

5' NNNNN $\underline{\text{nn}}$ cuga $\underline{\text{u}}$ g $\underline{\text{agg}}$ CCGTAAGG $\underline{\text{cc}}$ ga $\underline{\text{a}}$ a $\underline{\text{cc}}$ NNNNNN 3'

Sequence ID No: 48 Length: 28

5' NNNNN $\underline{\text{nn}}$ cuga $\underline{\text{u}}$ g $\underline{\text{ag}}$ pr pr pr pr $\underline{\text{c}}$ ga $\underline{\text{a}}$ a $\underline{\text{cc}}$ NNNNNN 3'

In addition to the methods of treating breast cancer or ovarian cancer using ribozymes of the invention, inhibitory compounds identified by the screening methods described previously can similarly be used to reduce the severity of breast cancer or ovarian cancer. Small organic compounds have particular advantages because of their ease of formulation and administration using well known methods in the pharmaceutical arts.

The present invention provides methods to identify gene targets that are down regulated by ribozyme activity. Another powerful application of this technology is to identify other genes in the pathway that contribute to the phenotype changes observed after ribozyme knockdown. Expression profiles can be determined in hybridization arrays of cDNA targets from known genes or expressed sequence tag (EST) sequence databases.

In one embodiment, differentially expressed mRNA in control ribozyme and effector ribozyme transduced cells, can be detected by reverse transcribing the mRNA from the cells into cDNA, which is labeled with green (control ribozyme) and red fluorescent dye (effector ribozyme). Individual EST clones are arrayed on a solid support such as polylysine-coated glass (*e.g.*, a 24 x 24 panel). The microarray is hybridized with a mix of the two labeled cDNA and then scanned for fluorescent color. A G3pDH DNA fragment is spotted on the corners of each 24x24 panel and serves as an internal control to check for an equal amount of cDNA applied from the two cell lines. The intensity of the color at each location of the array is proportional to the expression level of that gene in the sample. EST fragments which showed a predominant red color indicate that the EST is expressed mostly in the effector ribozyme cells or up-regulated by effector ribozyme treatment, while a predominance of green color indicates that the EST is expressed largely in the control ribozyme cells or down-regulated by effector ribozyme treatment.

In the above approach, arrays of clones can be formed on any of a variety of solid supports, including, for example a membranes such as made from nylon or a silicon based chip. Also, the cDNA prepared from the two transduced cell types may be differentially labeled by fluorophores, enzymes, radioisotopes and the like. If radiolabeled cDNA is used, each pool of labeled cDNA is hybridized to separate membranes, which are scanned for radioactive intensity using a phosphorimager. The radioactive intensity of each array element is proportional to the number of bound cDNA molecules, so the intensities of effector ribozyme cDNA can be directly compared to control ribozyme cDNA in a similar fashion to the fluorescent elements described above.

Criteria such as at least a 2 fold difference in expression level can be set to delineate potential targets. for genes involved in the pathway leading to the phenotypic change. Using this technology, additional drug targets indirectly regulated by the ribozyme can be identified.

Regulators of BRCA-1 expression were identified using the above expression profiling approach as described in greater detail in Example X. These genes, which included GenBank Acc. No. AA419229, GenBank Acc. No. H18950, GenBank Acc. No. H07920, GenBank Acc. No. H70047, GenBank Acc. No. H84815, GenBank

Acc. No. AA757764, GenBank Acc. No. H19111, GenBank Acc. No. AA629897, GenBank Acc. No. AA663439, GenBank Acc. No. R15740, GenBank Acc. No. AA454570, GenBank Acc. No. AA485748, GenBank Acc. No. AA464601, GenBank Acc. No. AA102107: and GenBank Acc. No. Z49826, represent therapeutic drug or ribozyme effector targets for increasing BRCA-1 expression.

The invention also provides a high throughput drug discovery method using ribozyme transduced cells and chip array technology. By combining array technology with ribozyme knockdown, drugs can be rapidly screened for effects on a given pathway. Once the expression profile leading to a given phenotype is determined, additional arrays can be generated with the relevant regulated EST sequences. These can be screened with mRNA from drug treated cells. Profiles matching the ribozyme treated profile can be identified. Treatment with the drugs identified in this way can be expected to give the desired phenotype.

This methodology allows the linking of the function of these target genes to the desired phenotype *i.e.*, increased BRCA-1 expression. Small molecule drugs, ribozyme drugs, or antibody drugs can be identified by those skilled in the art that inhibit the activity of these gene targets resulting, for example, in increased BRCA-1 expression. The gene targets can be used to develop high throughput assays that can be screened with existing small molecule libraries. In addition, genes which express a surface or secreted protein can be targets for antibody development. Antibodies specific for the gene product can be generated preferably in transgenic mouse systems to generate human antibodies. Furthermore, chimeric ribozyme drugs targeting these BRCA-1 regulators can be designed as explained above (see *e.g.*, scheme 1 above and scheme 2 below).

The invention also provides a method of detecting a neoplastic cell in a sample. In one embodiment, the method consists of contacting the sample with a detectable agent specific for a BRCA-1 regulator nucleic acid molecule, and detecting the nucleic acid molecule in the sample. Altered expression or structure of the nucleic acid molecule indicates the presence of a neoplastic cell in the sample. In another embodiment, the method consists of contacting the sample with a detectable agent specific for a BRCA-1 regulator polypeptide of the invention, and detecting the

polypeptide in the sample. Altered expression or structure of the polypeptide indicates the presence of a neoplastic cell in the sample.

The diagnostic methods described herein are applicable to the identification of neoplastic cells present. in solid tumors (carcinomas and sarcomas) such as, for example, breast cancer, ovarian cancer and prostate cancer. The BRCA-1 regulator molecules that can be detected using this method include nucleic acids encoding BR1, BR2, BR3, BBC1, or ID4 or the corresponding polypeptides.

Various qualitative and quantitative assays to detect altered expression or structure of a nucleic acid molecule in a sample are well known in the art, and generally involve hybridization of a detectable agent, such as a complementary primer or probe, to the nucleic acid molecule. Such assays include, for example, *in situ* hybridization, which can be used to detect altered chromosomal location of the nucleic acid molecule, altered gene copy number, or altered RNA abundance, depending on the format used. Other assays include, for example, RNA blots and RNase protection assays, which can be used to determine the abundance and integrity of RNA; DNA blots, which can be used to determine the copy number and integrity of DNA; SSCP analysis, which can detect single point mutations in DNA, such as in a PCR or RT-PCR product; and coupled PCR, transcription and translation assays, such as the Protein Truncation Test, in which a mutation in DNA is determined by an altered protein product on an electrophoresis gel. Further assays include methods known in the art for genotyping a BRCA-1 regulator gene, for example, by RFLP analysis or by determining specific SNPs in a BRCA-1 regulator gene. An appropriate assay format and detectable agent to detect an alteration in the expression or structure of a BRCA-1 regulator nucleic acid molecule can be determined by one skilled in the art depending on the alteration one wishes to identify.

Various assays to detect altered expression or structure of a BRCA-1 polypeptide are also well known in the art, and generally involve hybridization of a detectable agent, such as an antibody or selective binding agent, to the polypeptide in a sample. Such assays can be performed *in situ*, such as by immunohistochemistry or immunofluorescence, in which a detectably labeled antibody contacts a polypeptide in a cell. Other assays, for example, ELISA assays, immunoprecipitation, and

immunoblot analysis, can be performed with cell or tissue extracts. Assays in which the polypeptide remains in a native form are particularly useful if a conformation-specific binding agent is used, which can detect a polypeptide with an altered structure. A structural variant of a BRCA-1 regulator polypeptide can act, for example, in a dominant-negative fashion to decrease BRCA-1 expression or activity and cause unregulated cell proliferation. An appropriate assay format and detectable agent to detect an alteration in a BRCA-1 regulator polypeptide can be determined by one skilled in the art depending on the alteration one wishes to identify.

Also contemplated herein are *in vivo* methods for detecting a BRCA-1 polypeptide or nucleic acid, comprising administering to a subject a detectable agent such as an antibody or a complementary RNA molecule, appropriately conjugated with an imaging reagent useful for *in vivo* detection of the detectable compound using methods known in the art such as magnetic resonance imaging, computed tomography, x-ray imaging, and the like. Imaging reagents are well known in the art and include radioactive reagents, electron-dense reagents and magnetic or electronic reagents.

The diagnostic methods described herein can also be used in prognostic assays. Such an application can identify alterations in expression or structure of BRCA-1 regulator molecules that take place at characteristic stages in the progression of a proliferative disease or of a tumor. Knowledge of the stage of the tumor allows the clinician to select the most appropriate treatment for the tumor and to predict the likelihood of success of-that treatment. The diagnostic methods described herein can also be used to monitor the effectiveness of therapy. Successful therapy can be indicated, for example, by a reduction in the number of neoplastic cells in an individual, as evidenced by more normal expression or structure of the BRCA-1 regulator in a sample following treatment.

In the diagnostic and prognostic assays described herein, the level of expression or structure of the detected BRCA-1 regulator nucleic acid or polypeptide in the test sample is compared to the known expression level or structure of the nucleic acid or polypeptide in a normal sample. The normal sample can be obtained either from normal tissue of the same histological origin of the same or a different individual, or a population of individuals.

The diagnostic methods described herein can also be used in a method for determining the propensity of a subject to develop cancer. Such a method includes collecting a pre-neoplastic sample from a subject which can include a sample of a specific tissue or organ in which a neoplastic growth is suspected to occur, or alternatively a sample from any portion of the subject, such as a blood, saliva or cheek swab sample. The detected expression level or structure of a BRCA-1 regulator nucleic acid or polypeptide is then compared to the corresponding expression level or structure known for a population of normal and cancer subjects, whereby the expression level or structure will correspond to a specific subpopulation of subjects. The number of normal subjects relative to cancer subjects in the specific subpopulation can be determined and results in a likelihood value of any subject within that subpopulation for developing cancer. The propensity of the subject to developing cancer is determined by assigning to it the likelihood value of the subpopulation to which the expression level or structure of the subject corresponds.

It is understood that modifications which do not substantially affect the activity of the various embodiments of this invention are also included within the description of the invention provided herein. Accordingly, the following examples are intended to illustrate but not limit the present invention.

EXAMPLES

EXAMPLE I:

Preparation of the random retroviral vector ribozyme library

This example demonstrates the construction of a random retroviral vector ribozyme gene library. The inventors have discovered that by introducing a random retroviral vector ribozyme gene library in the PA1 ovarian carcinoma cell line, certain of the ribozymes will selectively target and inactivate mRNA molecules encoding regulators of the BRCA-1 tumor suppressor gene. A ribozyme gene library with randomized target recognition sequences was introduced into mammalian cells stably expressing a selectable marker, enhanced green fluorescent protein (EGFP) under the control of the BRCA-1 promoter. Cells in which BRCA-1 expression was

upregulated by particular ribozymes were selected through their concomitant increase in EGFP expression. The ribozyme genes were then rescued from the cells with enhanced EGFP expression and sequenced across their substrate binding sites. The corresponding ribozyme binding sequence, or "target sequence tag" (TST) is a sequence present in the regulator nucleic acid molecule targeted by the ribozyme. Thus, knowledge of the TST allows novel regulator nucleic acids to be identified and isolated.

A plasmid-based retroviral library was constructed by inserting random ribozyme gene sequences into parent vector pLHPM-2kb. pLHPM-2kb contains 5' and 3' long terminal repeats (LTR) of the Moloney retroviral genome; a neomycin resistance gene driven by the LTR; an SV40 promoter driving a puromycin resistance gene; and a transcription cassette containing a tRNAval promoter and a 2 kb stuffer insert. When the stuffer insert is removed and replaced by the random ribozyme library inserts, the tRNAval promoter can drive transcription of the inserted ribozyme gene.

To prepare the pLHPM-2kb vector, plasmid pLHPM was digested overnight at 65°C with BstB1, phenol: chloroform extracted and ethanol precipitated. The resuspended DNA was then digested overnight at 37°C with MluI. This double digestion excises the 2kb stuffer fragment. The resultant 6kb plasmid vector DNA fragment was purified by agarose gel electrophoresis.

To prepare the random ribozyme library inserts, three oligonucleotides were synthesized and annealed in annealing buffer (50 mM NaCl, 10 mM Tris pH 7.5, 5 mM MgC12) at a molar ratio of 1:3:3 (oligol:oligo2:oligo3) by heating to 90°C followed by slow cooling to room temperature. The three oligonucleotides had the following sequences:

Oligol: 5'–pCGCGTACCAGGTAATATACCACGGACCGAAGTCCG
TGTGTTTCTCTGGTNNNNTTCTNNNNNNNNGGATCCTGTTTCCGCC
CGGTTT–3'
(SEQ ID NO: 49)

Oligo2: 5'–pGTCCGTGGTATATTACCTGGTA–3' (SEQ ID NO: 50)

Oligo3: 5'–pCGAAACCGGGCGGAAACAGG–3' (SEQ ID NO: 51)

To provide for random and uniform incorporation of A, T, C and G nucleotides at the positions represented as N nucleotides in oligo1, the A, T, C and G reagents were premixed, and the same mixture used for every N position in the oligonucleotide synthesis. The ribozyme insert library formed by annealing the three oligonucleotides (SEQ ID NOS: 7–9) thus contains 8 positions with random nucleotides corresponding to helix 1 of the ribozyme, and 4 random positions with random nucleotides corresponding to helix 2 of the ribozyme.

In order to ligate the pLHPM-2kb vector DNA fragment with the random ribozyme insert library, 0.5 pmole of the vector and an 8-fold molar excess of annealed oligonucleotides were ligated overnight with 10 units of T4 DNA ligase. Ultracompetent DH12S bacteria were then electroporated with the ligation mixture. A total of 5 x 10^7 bacterial colonies containing the retroviral plasmid ribozyme library was obtained.

The bacterial colonies containing the retroviral plasmid ribozyme library were pooled in aliquots as a master stock and frozen at -80°C. Working stocks were made by culturing 1 ml of the master stock in 60 ml LB media overnight at 30°C. A 1 ml aliquot of the working stock was used to make a 500 ml bacterial culture by incubation at 30°C overnight. Plasmid DNA was then-extracted from the 500 ml culture and used to generate retroviral vector as described below.

Following the cloning of the randomized hairpin ribozyme genes into pLHPM, the "randomness" of the plasmid library was evaluated by both statistical and functional analyses. A complete ribozyme library of this design, with 12 random positions, would contain 4^{12}, or 1.67 x 10^7, different members. For the statistical analysis, forty individual bacterial transformants were picked and sequenced. This allowed an evaluation of the complexity of the library, without having to manually sequence each library member. The statistical "randomness" of the library was determined utilizing the formula for a two-sided approximate binomial confidence interval: E= 1.96*squareroot(P* (1–P) /N), where P= the expected proportion of each nucleotide in a given position (this value for DNA bases equals 25% or P=0.25); E=

the desired confidence interval around P (i.e. P+/-E); and N= the required sample size (Callahan Associates, Inc., La Jolla, CA). To determine the proportion of each base within 5% (E=0.05), the required sample size is 289. Since each ribozyme molecule contains twelve independent positions, the number of individual ribozyme genes that need to be sequenced out of the library equals 289 divided by 12, or about 25 molecules.

The frequencies of the four nucleotides, with 95% confidence limits, in the random positions were calculated to be G: 22.3 +6.1, A: 31.9+7.0, T: 27.3+7.8 and C: 18.01+15.1. Since the expected frequency for each base is 25%, each base appears to be randomly represented (except for C, which may be slightly lower than expected). These variations most likely result from the unbalanced incorporation of nucleotides during the chemical synthesis of the oligonucleotides and could reduce the complexity of the library.

For a functional evaluation of the library's complexity, *in vitro* cleavage was utilized to determine if ribozymes that target known RNA substrates were present in the library pool. This involved *in vitro* transcribing of the entire ribozyme library in one reaction and then testing the pool's ability to cleave a variety of different RNA substrates of both cellular and viral origin. Six out of seven known RNA targets were properly and efficiently cleaved by the *in vitro* transcribed library. This qualitative analysis suggested a significantly complex library of ribozyme genes and the lack of cleavage of one target out of seven may reflect the slight non-randomness suggested by the base composition described above.

Viral vector was produced from the ribozyme library plasmid using a triple transfection technique. CF2 cells were seeded at 3.5 x 10^4 cells/cm^2 one day prior to transfection. The cells were transfected with a 1:1:1mixture of the ribozyme library plasmid or control ribozyme plasmid, a plasmid encoding the moloney-murine virus gag-pol genes, and a plasmid encoding the vesicular stomatitis virus-G gene, using the cationic lipid TransIT-LT1 (Pan Vera Corporation). 7.8 x 10^6 cells were transfected with 25 mg of each plasmid complexed with 250 ml of the lipid in a total volume of 20 ml of serum free media. After 6 hours, the media was replaced with growth media. The cell supernatant containing retroviral particles was collected every 24 hours beginning on day 2 after addition of fresh media. The supernatant

was filtered through 0.4 mm filters and titered in a standard assay using HT1080 cells.

EXAMPLE II:

Isolation of ribozymes that target BRCA-1 regulator nucleic acids

This example demonstrates the isolation of ribozyme genes that bind to and inactivate BRCA-1 regulator nucleic acid molecules, and the identification of the nucleic acid sequences they target.

A reporter gene based cellular selection system was used in order to facilitate functional selection for cells with a change in their BRCA-1 expression. The reporter construct allowed expression of EGFP under control of the BRCA-1 promoter region contained in a 2.9 kb fragment from the BRCA-1 genomic sequence. The 2.9 kb BRCA-1 promoter region was amplified using a primer with an added PstI site (5'–ATCTTTCTGCAGCTGCTGGCCCGG–3'; SEQ ID NO: 52) and an AgeI containing primer (5'–GTGTAAACCGGTAACGCGAAGAGCAGATA–3'; SEQ ID NO: 53) and the PCR Long Template System (Boehringher Mannheim) from a pGL2 vector containing a 3.8 kb genomic BRCA-1 fragment. The PCR product was gel purified, digested with PstI and AgeI and cloned into pEGFP-1 (Clontech) which also contains a neomycin resistance marker using standard molecular techniques.

The ovarian cancer cell line PA1 was transfected with the BRCA/EGFP reporter construct using Lipofectamine Plus (Gibco-BRL) and stable integrants selected using neomycin at 400-800 mg/ml for two weeks. Clones of reporter cells were generated by limited dilution and characterized for endogenous BRCA-1 and EGFP expression. A reporter clone designated P8EGBR3 with an intermediate level of expression of both genes was chosen for library introduction. Retroviral library vector or control vector was introduced into the P8EGBR3 reporter cells, and selected for stable integration of the vector using the puromycin resistance marker. 1. 9×10^7 cells were transduced with retroviral vector (library or control) at an MOI of 1 based on the HT1080 titer. Cells were split at a 1:2 ratio two days after transduction and refed with media containing 0.3 mg/ml puromycin for 14 days to select cells with stable integration of the retroviral vector.

After stable selection, the cell populations were subjected to several rounds of fluorescence activated cell sorting (FACS), enriching for cells with an increase in EGFP expression (the highest three to ten percent of respective populations). After three rounds of selection, the ribozyme library transduced population revealed a subpopulation with an increase in EGFP expression, while the control transduced cells did not. After a further round of selection, the majority of the library transduced cells had shifted (FIGURE 3), with a five-fold increase in mean fluorescence intensity (MFI) compared to control vector transduced cells (FIGURE 4). RNA analysis from these populations revealed an increase of EGFP message as well as endogenous BRCA-1 message in the library populations compared to control vector transduced cells (FIGURE 5). These data underline the ability of selected ribozymes to increase the reporter gene expression concomitant with endogenous BRCA-1 expression, indicating a regulation on the transcriptional level.

Ribozymes were rescued from the shifted cell population by PCR amplification. PCR rescue was performed on five separate aliquots of 1 mg of genomic DNA extracted from the cells using the QIAmp Blood Kit (Qiagen, Valencia, CA). PCR was carried out using the AmpliTaq Gold system (Perkin-Elmer, Norwalk, CT) with an initial denaturation at 94°C for 10 min. followed by 35 cycles of 94°C for 20 sec., 65°C for 30 sec., and 72°C for 30 sec. A final extension was performed at 72°C for 7 min. PCR primers (5'–GGCGGGACTATGGTTGCTGACTAAT–3', 5'–GGTTATCACGTTCGCCTCACACGC–3'; SEQ ID NOS: 54 and 55, respectively) within the vector amplified a 300 bp fragment containing the ribozyme genes. The pooled PCR product, which contained a pool of ribozyme genes, was isolated by electrophoresis on 1% agarose, purified using a Gel Extraction Kit (Qiagen), then digested with BamHI and MluI and ligated into LHPM digested with the same enzymes. The ligated DNA was used to transform DH12S bacteria by electroporation. The entire bacterial culture was plated on LB-agar plates containing ampicillin and incubated at 37°C overnight. The resulting bacterial colonies were pooled and purified DNA was used in the triple transfection protocol to generate new retroviral vector. Individual colonies were also sequenced by the standard dideoxy method using a vector primer (5'–CTGACTCCATCGAGCCAGTGTAGAG–3';

SEQ ID NO: 98). Several ribozyme sequences were found to be enriched. The retroviral pool was introduced into new reporter cells and subjected to several rounds of selection by FACS. Again the library transduced cells shifted compared to control transduced cells. Rescue of the ribozymes from these cells showed six predominant ribozymes designated. RH1, RH2, RH3, RH4, RH5, and RH6 (Tables 1–6).

The substrate binding sequence of RH1 (SEQ ID NO: 5), together with its corresponding target sequence tag (TST1; SEQ ID NO: 11) is presented in Table 1 below.

Table 1

RH1 gene sequence	Corresponding TST1
CCGGATGC AGAA CAAT (SEQ ID NO: 5)	ATTG NGTC GCATCCGG (SEQ ID NO: 11)

The substrate binding sequence of RH2 (SEQ ID NO: 10), together with its corresponding target sequence tag (TST2; SEQ ID NO: 16) is presented in Table 2 below.

Table 2

RH2 gene sequence	Corresponding TST2
CCCTATTT- AGAA TTGT (SEQ ID NO: 10)	ACAA NGTC AAATAGGG (SEQ ID NO: 16)

The substrate binding sequence of RH3 (SEQ ID NO: 6), together with its corresponding target sequence tag (TST3; SEQ ID NO: 12) is presented in Table 3 below.

Table 3

RH3 gene sequence	Corresponding TST3
AGTACATT AGAA TACT (SEQ ID NO: 6)	AGTA NGTC AATGTACT (SEQ ID NO: 12)

The substrate binding sequence of RH4 (SEQ ID NO: 7), together with its corresponding target sequence tag (TST4; SEQ ID NO: 13) is presented in Table 4 below.

Table 4

RH4 gene sequence	Corresponding TST4
CTAGTGAG AGAA GGGA (SEQ ID NO: 7)	TCCC NGTC CTCACTAG (SEQ ID NO: 13)

The substrate binding sequence of RH5 (SEQ ID NO: 8), together with its corresponding target sequence tag (TST5; SEQ ID NO: 14) is presented in Table 5 below.

Table 5

RH5 gene sequence	Corresponding TST5
TGAGATCC AGAA AAGC (SEQ ID NO: 8)	GCTT NGTC GGATCTCA (SEQ ID NO: 14)

The substrate binding sequence of RH6 (SEQ ID NO: 9), together with its corresponding target sequence tag (TST6; SEQ ID NO: 15) is presented in Table 6 below.

Table 6

RH6 gene sequence	Corresponding TST6
TGTTACT AGAA TGTT (SEQ ID NO: 9)	AACA NGTC AGTAACA (SEQ ID NO: 15)

To verify the function of the predominating ribozymes, vector was generated from plasmids encoding individual ribozymes and introduced by retroviral transduction into a second reporter system (SKHYTK4) with. the hygromycin resistance gene driven by the BRCA-1 promoter in the SKBR3 breast carcinoma cell line. The BRCA-1/EGFP plasmid was used to excise a SapI insert containing the 2.9

kb BRCA-1 promoter region, the ends were blunted using Klenow and the insert digested with BglII. A plasmid containing the hygromycin resistance gene fused to HSV thymidine kinase was used as a backbone 30 to construct the hygromycin reporter plasmid. This plasmid contains a separate neomycin resistance cassette. The plasmid was digested with NheI, the ends were blunted using Klenow and the plasmid digested with BglII. The insert was ligated into the plasmid with one blunt end and one BglII site to generate a BRCAl promoter driven hygromycin gene. The breast cancer cell line SKBR3 was transfected with the BRCA/HyTK reporter construct using Lipofectamine Plus (Gibco-BRL) and stable integrants selected using neomycin at 400-800 mg/ml for two weeks. Clones of reporter cells were generated by limited dilution and characterized for resistance to hygromycin B. One clone, SKHYTK4, was selected for its sensitivity to hygromycin over a range of 400-800 mg/ml. The BRCA-1/hygroTK cassette was determined to be present in these cells by PCR and sequence analysis of the genomic DNA. SKHYTK4 cells transduced with individual ribozymes RH3, RH4, or RH6 showed increased resistance to hygromycin as determined by an increase in cell number when the transduced cells were grown in hygromycin containing media (3.1, 8.5, and 147 fold respectively). RH2 did not confirm in this assay.

In addition, individual ribozymes were introduced into PA1 cells by retroviral transduction, and analyzed for their ability to grow in soft agar. 6 well plates were coated with Iscove's media (Gibco-BRL) containing 0.6% agar and 10% fetal calf serum. 1 $x10^6$ PA1 cells were transduced with retroviral vector at an MOI of 1 on day I and 2. On day 3, freshly transduced PA1 cells were overlaid on coated plates at 6×10^4 or 9×10^4 cells per well in Iscove's media containing 0.4% agar and 7.7% fetal calf serum. On day.4, an upper layer of Iscove's media containing 0.6% agar and 10% fetal calf serum was applied. Every third day, Iscove's media containing 10% fetal calf serum was replaced on the surface of the agar layer. The cells were incubated for 6 weeks at 37°C, 5% $C0_2$ and the number of colonies greater than 100 cells in size was determined. PA1 cells transduced with individual ribozymes RH1, 4, or 5 showed a decreased ability to grow in soft agar, indicating upregulation of the endogenous BRCA-1 message leading to suppression of anchorage-independent growth (3.7, 26, and 28% of control respectively). RH2 did not confirm in this assay.

In view of their ability to reproducibly cause an increase in expression from the BRCA-1 tumor suppressor promoter and to induce changes in the ability of PA1 cells to grow in soft agar, ribozymes containing substrate binding sequences designated SEQ ID NOS: 5–10 are ribozymes which target and inactivate regulators of BRCA-1 expression nucleic acid molecules. Likewise, the targets of these ribozymes, which are nucleic acid molecules containing nucleic acid sequences designated SEQ ID NOS: 11–16, are regulators of BRCA-1 expression nucleic acid molecules.

EXAMPLE III

Isolation and characterization of Breast Basic Conserved Protein 1 (BBC1)

This example demonstrates the isolation of a full-length transcriptional regulator of BRCA-1 nucleic acid molecule designated Breast Basic Conserved Protein 1 cDNA and its encoded polypeptide.

Since ribozymes recognize their cognate targets by sequence complementarity, the sequence of a ribozyme that causes a phenotype through its catalytic activity predicts a sequence tag that can be used to clone the target gene. This "Target Sequence Tag" or TST is about 16 bases long, consisting of the two target regions complementary to ribozyme helices 1 and 2 and the requisite GUC (see, for example Figures 1, 2 and 7). The TST can thus be used to BLAST search the gene and EST databases, and also can be used as a primer for 3' and 5' RACE. BLAST searches of the databases with the TST corresponding to RH1 yielded no complete matches, and incomplete matches only with non-human sequences.

In light of the absence of obvious database candidates, the RH1 target gene was cloned using the RK TST as a primer for a novel technique developed for the purpose of identifying ribozyme cleavage targets. This novel technique called SMART C-SPACE for Switching Mechanism At 5' end of RNA Template in Cleavage-Specific Amplification of cDNA Ends takes advantage of the unique cleavage site of the hairpin ribozyme to generate a 5' terminus on the target nucleic acid consisting of the bases GUC.

This method consists of *in vitro* cleavage of cellular mRNA by a particular ribozyme (compared with a control ribozyme cleavage performed with another aliquot of the target cell mRNA), combined with SMART cDNA technology (Clontech) for detection of the cleavage products by PCR amplification. SMART cDNA technology is a PCR-based method for generating high yields of high quality, double-stranded cDNA from nanograms of total or poly A+ RNA. Poly A+ RNA from PA1 cells was first cleaved with *in vitro* transcribed RH1 ribozyme (Welch P et al., 1997), then used as a template for SMART cDNA synthesis. The first strand reaction was primed by a modified oligo (dT) oligonucleotide and catalyzed by MMLV reverse transcriptase (Superscript, GIBCO BRL). When the enzyme reaches the 5' end of the cleaved (or uncleaved) mRNA, the enzyme's terminal transferase activity adds 3-5 deoxycytidine nucleotides to the 3' end of the first strand cDNA. The 3' end of the SMART oligonucleotide, which contains a stretch of G residues, anneals with the C-rich motif, forming an extended template. The reverse transcriptase enzyme then switches templates and replicates the oligonucleotide. After RNase treatment, the result is a single-stranded cDNA with a sequence complementary to the oligo (dT) primer at the 5' end and a sequence complementary to the SMART oligonucleotide at the 3' end. These sequences subsequently serve as priming sites in PCR amplification. In the case of ribozyme cleaved RNA fragments, the 3' attached sequence lies directly adjacent to the CAG sequence generated by reverse transcription of the ribozyme recognition triplet GUC of the mRNA. This sequence stretch, in combination with the adjacent 8 bases of helix 1 of the ribozyme consists of a stretch of 14 bases specific for cleaved substrate RNA molecules. The cDNA generated using this technique was amplified by PCR using an oligonucleotide primer that recognize the SMART sequence, attached C residues, GTC site, and adjacent helix 1 of the ribozyme, and an oligonucleotide primer that recognizes a unique portion of the oligo(dT) primer. The amplification products were compared between the RH1 cleaved mRNA and mRNA cleaved with a control ribozyme. Several bands specific for the RH1 ribozyme cleaved sample were recovered, cloned and sequenced (Figure 6).

The sequence of one of the SMART C-SPACE bands matched that of the Breast Basic Conserved Protein 1(GenBank accession number X64707), with a

13/16 match at the target recognition site (Scheme 2). This gene was originally identified by differential screening of a human breast carcinoma cDNA library. It has been found to be more abundant in fibroadenomas than in carcinomas. It encodes two potential nuclear localization signals. The protein encoded by this gene contains a 25 amino acid region which exhibits strong similarity with the plant basic peptide P14, which represents a new class of transcriptional activators. It is localized on chromosome 16q24.3. The BBC1 message is found to be downregulated in hormone refractory prostate cancer.

<u>Scheme 2. Comparison of RH1 TST sequence (SEQ ID NO: 11) and BBC1 sequence (SEQ ID NO: 1)</u>

RHI		5'–ATTG NGTC GCATCCGG–3'
BBC1	nt 208-223	5'–GTCG GGTC CCATCCGG–3'

To confirm the involvement of BBC1 in the transcriptional regulation of BRCA-1, additional ribozyme target GUC sites were identified in the BBC1 sequence, and target validation ribozymes designed (TV6-10, Table 7; SEQ ID NOS: 56 through 60, respectively). Target validation ribozyme expressing retroviral vectors generated as before from the individual target validation ribozyme sequences were tested as described above in the SKHYTK4 hygromycin resistance assay. Several of these target validation ribozymes scored positive for enhanced hygromycin resistance with TV9 showing the most pronounced effect (20 fold increase in cell number). Additional target validation ribozymes are provided in Figure 8 (SEQ ID NOS: 102 through 221).

Table 7

Validation ribozyme	Gene sequence (SEQ ID NO)
TV6	GGCTTCAA AGAA ATGC (16)
TV7	TGGGACCC AGAA CGGG (17)
TV8	CCGGATGG AGAA CGAC (18)
TV9	ACGTTCCG AGAA GGCA (19)
TV10	CCCAGCAT AGAA GCCC (20)

To directly measure the increased expression of BRCA-1 resulting from the inhibition of the transcriptional regulator BBC1, real-time PCR was performed on RNA from cells with and without the inhibitory ribozyme. Real-time quantitative polymerase chain reaction (RT-PC-R) is a process by which the cellular levels of mRNA encoding various proteins can be determined. In order to perform RT-PCR three oligonucleotides based on the sequence of the mRNA of interest and a special thermal cycler with a fluorescence detector are needed. Two of these oligonucleotides act as primers just as in regular PCR. The third oligonucleotide acts as a "probe" and must match sequence which lies between the two primers. This probe has two groups attached. A reporter dye is attached to the 5' end and a quencher is attached to the 3' end. When these two groups are close together, as they are when attached to the two ends of the oligos, the fluorescence from the reporter dye is quenched by the quencher. In this arrangement, the fluorescence detector will not be able to detect any signal. However, during PCR, the probe anneals to its template between the two primers. As the taq polymerase enzyme replicates the template from the primers, it comes upon the probe. Besides its polymerase activity, taq also has a 5' to 3' exonuclease activity. Therefore, when it reaches the place in the template where the probe is annealed, it will displace the probe and degrade it from its 5' end. This releases the reporter dye from the oligonucleotide, and thus allows the reporter dye to diffuse away from the quencher. Once the reporter dye is separated from the quencher, the fluorescence detector can detect the signal. This fluorescence reading is taken at the end of each cycle of the PCR. In this way, the detector can determine the amount of probe which has been destroyed during each cycle. The

amount of probe destroyed is directly proportional to the amount of template at the beginning of the PCR reaction. Consequently, the amount of mRNA encoding a given protein can be determined in each sample tested. For determining BRCA-1 mRNA levels, we use the following oligonucleotide sequences:

forward primer– 5'–CTGCTCAGGGCTATCCTCTCA–3'

reverse primer– 5-TGCTGGAGCTTTATCAGGTTATGT-3

probe– 5'–1TGACATTTTAACCACTCAGCAGAGGGATACCA2–3'

where 1 is the 6-FAM reporter dye and 2 is the TAMRA quencher (SEQ ID NOS: 61 through 63, respectively). RT-PCR-analysis of the mRNA extracted from the TV9 transduced cells showed a 150% increase in BRCA-1 message compared to control vector transduced cells. This confirms the involvement of BBC1 in the observed phenotype, and establishes a novel function of this gene as a transcriptional regulator of BRCA-1.

An additional SMART C-SPACE product was identified with a 12/16 match to CHLR2 (GenBank accession number U33834), a partial sequence-found in the database purported to be a helicase (Scheme 3). Validation of this sequence with 5 additional ribozymes (TV 11-15, Table 8; SEQ ID NOS: 65 through 69) targeting GUC sites failed to show enhanced hygromycin resistance. This demonstrates the requirement for the validation technology described in this application in order to distinguish between several possible candidate genes.

Scheme 3. Comparison of RH1 TST sequence (SEQ ID NO: 11) and CHLR2 sequence (SEQ ID NO: 64)

RH1		5'–ATTG NGTC GCATCCGG–3'
CHLR2	nt 1621-1636	5'–GGTG GGTC GCATCCTC–3'

Table 8

Validation ribozyme	Gene sequence (SEQ ID NO)
TV11	CCGAGAGA AGAA AGCC (65)
TV12	TGGTTGGA AGAA CCGA (66)
TV13	GAGGATGC AGAA CCAC (67)
TV14	AAGAAACA AGAA ACCC (68)
TV15	TT.GGCCAG AGAA GGGG (69)

EXAMPLE IV

Isolation and characterization of Candidate BRCA-1 Regulator (BR1)

This example demonstrates the isolation of a full-length regulator of BRCA-1 nucleic acid molecule designated Candidate BRCAl Regulator (BR1) cDNA and its encoded polypeptide.

BLAST searches of the E8T databases with RH4 identified a 15/16 match with an EST (GenBank accession number AA886839) without similarity to a known gene. The RH4 TST and sequences from the identified EST were used as primers for 5' and 3' RACE to clone a 2.9 kb cDNA (Figure 7). This novel sequence was designated Candidate BRCA-1 Regulator (BR1). The sequence has an ORF of 127 amino acids with significant similarity to elongation factor g. However the C terminal domain of the ORF is not conserved. This can indicate a function for this gene in the regulation of translation of the BRCA-1 message. Interestingly, the PTI-1oncogene is a truncated elongation factor la.

To confirm the involvement of BR1 in the regulation of BRCA-1, additional ribozyme target GUC sites were identified in the BR1 sequence, and target validation ribozymes designed (TV1–5, Table 9; SEQ ID NOS: 70 through 74, respectively). Retroviral vectors generated as before from the individual target validation ribozyme sequences were tested in the SKHYTK4 hygromycin resistance assay. Several of these target validation ribozymes scored positive for enhanced hygromycin resistance, with TV3 showing a 17 fold increase in cell number. This confirms the involvement of BR1 in the observed phenotype, and establishes a novel function of this gene as a regulator of BRCA-1.

Table 9

Validation ribozyme	Gene sequence (SEQ ID NO)
TV1	TTGATGTG AGAA GCTT (70)
TV2	ACTTTTCT AGAA GGAA (71)
TV3	TATTCCAT AGAA ACTG (72)
TV4	AGGACTGG AGAA AGCC (73)
TV5	AACACATT AGAA TCAA (74)

EXAMPLE V

Isolation and characterization of Inhibitor, dominant negative 4 (ID4).

This example demonstrates the isolation of a full-length transcriptional regulator of BRCA-1 nucleic acid designated Inhibitor, dominant negative 4 (ID4) cDNA and its encoded polypeptide.

The RH5 TST sequence showed a 14/16 match to the ID4 gene (GenBank accession number NM_001546), by BLAST search. ID4 is a known transcriptional repressor located on chromosome 6p21.3-p22/6p22.3-p23. It belongs to a family of inhibitors of basic helix-loop-helix transcription factors. It prevents transcriptional activation by dimerization to the factors and blocks their ability to bind DNA. They are known to be involved in differentiation of adipocytes and neuronal cells.

To verify the involvement of ID4 in the regulation of BRCA-1, target validation ribozymes (TV19–22, Table 10; SEQ ID NOS: 75 through 78, respectively) were designed corresponding to GUC sites in the ID4 sequence. Retroviral vectors generated from the individual target validation ribozymes were introduced into the SKHYTK4 cell line to assay for hygromycin resistance. The validation ribozymes as well as the original RH5 ribozyme failed to generate enhanced resistance to hygromycin. ID4 message was found to be undetectable in SKBR3 cells. The validation ribozymes were tested in the soft agar assay using PA1 cells, which do express ID4. RH5 and several target validation ribozymes decreased the anchorage-independent growth ability of these cells (58,10 and 34% reduction for

RH5, TV20, and TV21 respectively). Over-expression of the entire cDNA for ID4 in PAl cells caused an increase in the anchorage-independent growth ability of these cells, leading to large colonies after 10 days of culture. RT-PCR analysis of cells transduced with RH5 vector showed considerable variability depending on the population tested. This can be due to a complex feedback regulation mechanism.

Table 10

Validation ribozyme	Gene sequence (SEQ ID NO)
TV19	CAGTGGGC AGAA CTCA (75)
TV20	CAACAATT AGAA GGAG (76)
TV21	CACACCTG AGAA GCGC (77)
TV22	CGCGGCTG AGAA GGTC (78)

An additional possible candidate sequence was identified with a 13/16 match to AF6 , a translocation breakpoint sequence found in AML (GenBank accession number NM 005936). Validation of this sequence with 3 additional ribozymes (TV23, 26, 27, Table 11; SEQ ID NOS: 79 through 81) targeting GUC sites failed to show enhanced hygromycin resistance or growth in soft agar. This demonstrates the requirement for the validation technology described in this application in order to distinguish between several possible candidate genes.

Table 11

Validation ribozyme	Gene sequence (SEQ ID NO)
TV23	GTACTAGA AGAA CGAA (79)
TV26	TGTGATCC AGAA AAGG (80)
TV27	GGTGGCCA AGAA GTGG. (81)

EXAMPLE VI

Isolation and characterization of an EST corresponding to RH3

This example demonstrates the isolation of an EST encoding a nucleic acid containing a TST corresponding to RH3.

The RH3 TST sequence was used as a primer for 5' RACE using standard methodology. The resulting PCR product had a 13/16 match with an EST from the databases (GenBank accession number AL045940). This EST is undergoing target validation with 5 additional GUC recognition sites (SEQ ID NOS: 82 through 86, respectively).

Table 12

Validation ribozyme	Gene sequence (SEQ ID NO)
TV28	AAAAATTA AGAA GTCA (82)
TV29	GCTGTCCT AGAA TCAA (83)
TV30	TGTCAAAG AGAA CACC (84)
TV31	TGCAATGA AGAA ACTG (85)
TV32	TTACAATA AGAA ACTT (86)

A second EST was identified by database search (GenBank accession number AI668913) with a 14/15 match to RH3. This EST is undergoing target validation with 4 additional GUC recognition sites (SEQ ID NOS: 87 through 91, respectively).

Table 13

Validation ribozyme	Gene sequence (SEQ ID NO)
TV33	TCCTTCCAAGAAACGG (87)
TV34	TTGGCCCTAGAATGAG (88)
TV35	TTTTTCTAAGAAGCCC (89)
TV36	TTGTCTCAGAATGCG (90)
TV37	ATCGTCAAAGAAATCA (91)

EXAMPLE VII

Isolation and characterization of an EST corresponding to RH6

An EST was identified in the databases with a 15/15 match to RH6 (GenBank accession no. AI276397). Target validation ribozymes generated to 2 available GUC sites in this sequence (SEQ ID NOS: 92 and 93, respectively) failed to confirm in either the hygromycin resistance or soft agar assays indicating that an as yet unidentified gene is the real target responsible for the observed phenotype.

Table 14

Validation ribozyme	Gene sequence (SEQ ID NO)
TV16	CTATTTAA AGAA AATT (92)
TV17	TATTTCTT AGAA GTTC (93)

EXAMPLE VIII

Expression of BBC1, BR1, and ID4 in human cell lines

Northern analysis of mRNA extracted from the human ovarian carcinoma cell line PA1, and the breast carcinoma cell lines MCF7 and SKBR3 was done to determine the expression pattern of the genes. identified as BRCA-1 regulators. Total RNA was extracted from cell cultures grown at 80% confluency using the RNAeasy kit (Qiagen). Fifteen micrograms of total RNA was denatured in 50% formamide, 17 mM MOPS, 2.2 M formaldehyde, for 15 min. at 70°C. Samples were separated by electrophoresis in a 0.8% agarose gel containing 2.2 M formaldehyde and 20 mM MOPS and transferred to a nylon membrane for hybridization with 32P labeled probes generated from plasmids encoding either BBC1, BR1, or ID4. BBC1 and BR2 were found to be expressed in both PA1 and SKBR3 cells (MCF7 not done). ID4 was found to be expressed in both PA1 and MCF7 but not SKBR3 cells. To confirm this result, RT-PCR analysis using total RNA was performed as this is a more sensitive technique. Again, ID4 was found. to be expressed in both PA1 and

MCF7 but not SKBR3 cells. This explains the lack of effect of the ID4 targeted ribozymes observed in the SKBR3 based hygromycin resistance assay.

EXAMPLE IX

Effect of over-expression of ID4

To validate the biological relevance of ID4 regulation of BRCA-1, ID4 sense and antisense expression vectors were introduced into P8R3 cells by transduction with retroviral vectors. The cDNA sequence from ID4 was amplified using primers ID4-3 (5'–TCCGAAGGGAGTGACTAGGACACCC–3'; SEQ ID NO: 94) and ID4-7 (5'–TTCTGCTCTTCCCCCTCCCTCTCTA–3'; SEQ ID NO: 95). The amplified product was cloned into T/A-vector (Invitrogen), and verified by sequencing from both vector sites. Plasmid DNA was digested with Eco RI, releasing the whole insert out of the cloning vector and cloned into the EcoRI linearized retroviral expression vector pLPCX (Clontech). Clones were analyzed for the correct insert as well as for orientation screening by multiple restriction digests. Sequencing of both junction sites from both directions using vector-based primers 5 - and 3 - LPCX (Clontech) confirmed the correct orientation and junction sequence. Following stable integration, cells were evaluated for their anchorage-independent growth potential using the soft agar assay described above. The biological consequences of BRCA-1 downregulation should be increased growth in soft agar, and the converse for BRCA-1 upregulation. Cells expressing sense ID4 had a dramatic increase in anchorage-independent growth in comparison to cells expressing antisense ID4.

EXAMPLE X

Detection of genes in the Regulation Pathway of BRCA-1

This example provides methods to use BRCA-1 regulator ribozymes to identify other genes in the BRCA-1 pathway that contribute to the phenotype changes observed after ribozyme mediated target RNA cleavage (*i.e.* target "knockdown"). Expression profiling of genes in hybridization assays identifies

cellular genes upregulated or downregulated subsequent to BRCA-1 regulator knockdown. Such genes constitute potential BRCA-1 regulators.

RNA extracted from cells expressing ribozymes RH1, RH3, RH4, RH5, and RH6 was sequentially hybridized to an array filter. Probes, filters, and analysis were as recommended by the manufacturer (Research Genetics, Huntsville AL). Briefly, 10 micrograms of total RNA from each ribozyme or control transduced cell population was primed with oligo dT, and reverse transcribed with reverse transcriptase (Superscript II Life Technologies, Gaithersburg MD) incorporating ^{33}P labeled dCTP to generate a detectable cDNA probe. Denatured probe was hybridized to the filter in the supplied buffer for 18 hours at 42°C. Filters were washed in 2X SSC, 1% SDS at 50°C twice for 20 minutes each, followed by washing in 0.5 X SSC, 1% SDS at room temperature for 15 minutes. The filters were exposed to a phospor imaging screen and the image captured using a Molecular Dynamics (Sunnyvale, CA) phosphor imaging system. Comparisons between filters were made using the Pathways™ analysis software (Research Genetics, Huntsville AL).

The expression profile was compared to the pattern generated by hybridizing the filter to RNA extracted from cells expressing the control ribozyme. Several genes were found to be down-regulated in some or all of the ribozyme expression profiles. From the pattern of expression it is possible to ascertain a pathway of differentially expressed genes. Numerous genes were differentially expressed at levels in the ribozyme expression cells that were greater than in the control cells.

1. AA419229: EST sharing sequence identity with orphan GPR39
   A. AF034633.1 - Homo sapiens orphan G protein-coupled receptor (GPR39) mRNA, complete cds
   B. 4504096 - Homo sapiens G protein-coupled receptor 39 (GPR39) mRNA, and translated products
   C. Protein: O43194 – Human putative G protein-coupled receptor 39
   D. Protein: 4504097 - G protein-coupled receptor 39

2. H18950: EST sharing sequence identity with NADC3

A. AF154121.1 - Homo sapiens sodium-dependent high-affinity dicarboxylate transporter (NADC3) mRNA, complete cds

3. H07920: EST sharing sequence identity with MKK6
   A. XM_012648.1 - Homo sapiens mitogen-activated protein kinase kinase 6 (MAP2K6), mRNA
   B. NM_002758.1 - Homo sapiens mitogen-activated protein kinase kinase 6 (MAP2K6), mRNA
   C. U39657.1 - Human MAP kinase kinase 6 (MKK6) mRNA, complete cds

4. H70047: EST sharing sequence identity with RGS13
   A. AF030107 – Homo sapiens regulator of G protein signaling (RGS13) mRNA, complete cds

5. H84815: EST sharing sequence identity with Rab9 effector p40
   A. BC000503 - Homo sapiens, Rab9 effector p40, clone MGC:8459, mRNA, complete cds
   B. NM_005833.1 - Homo sapiens Rab9 effector p40 (RAB9P40), mRNA
   C. Z97074.1 - Homo sapiens mRNA for Rab9 effector p40, complete cds

6. AA757764: EST sharing sequence identity with multi Zn-finger TF
   A. D49835 – Homo sapiens mRNA for DNA-binding protein, complete cds

7. H19111: EST sharing sequence identity with sialyltransferase

8. AA629897: EST sharing sequence identity with laminin receptor
   A. M14199.1 - Human laminin receptor (2H5 epitope) mRNA, 5' end
   B. BC002533 - Homo sapiens, laminin receptor 1 (67kD, ribosomal protein SA), clone MGC:2208, mRNA, complete cds
   C. NM_002295.2 - Homo sapiens laminin receptor 1 (67kD, ribosomal protein SA) (LAMR1), mRNA

9. AA663439: EST sharing sequence identity with adenine nucleotide translocator 3
   A. J03592.1 - Human ADP/ATP translocase mRNA, 3' end, clone pHAT8
   B. AF076617 - Homo sapiens ADP/ATP translocase mRNA, 3'UTR
   C. L29034.1 - Human mitochondrial (clone GC6-T5) ADP/ATP translocase gene sequence

10. R15740: EST sharing sequence identity with chondroitin-6-sulfotransferase
   A. XM_012024.1 - Homo sapiens carbohydrate (chondroitin 6/keratan) sulfotransferase 1 (CHST1), mRNA
   B. NM_003654.1 - Homo sapiens carbohydrate (chondroitin 6/keratan) sulfotransferase 1 (CHST1), mRNA
   C. U65637.1 - Homo sapiens chondroitin-6-sulfotransferase mRNA, complete cds

11. AA454570: EST sharing sequence identity with lambda 5 semaphorin
   A. U38276.1 - Human semaphorin III family homolog mRNA, complete cds

12. AA485748: EST sharing sequence identity with fibromodulin
   A. XM_001782.2 - Homo sapiens fibromodulin (FMOD), mRNA
   B. NM_002023.2 - Homo sapiens fibromodulin (FMOD), mRNA

13. AA464601: EST sharing sequence identity with tspan-5
   A. AF065389 - Homo sapiens tetraspan NET-4 mRNA, complete cds
   B. AF053455 - Homo sapiens tetraspan TM4SF (TSPAN-5) gene, complete cds
   C. NM_005723.1 - Homo sapiens tetraspan 5 (TSPAN-5), mRNA
   D. XM_011178.1 - Homo sapiens sharing sequence identity with tetraspan 5 (H. sapiens) (LOC65435), mRNA

14. AA102107: EST sharing sequence identity with aminopeptidase A
   A. L12468.1 - Homo sapiens aminopeptidase A mRNA, complete cds

15. Z49826: HSHNF4G hepatocyte nuclear factor 4 gamma

Ribozymes can be designed to target the cleavage sites in the above 15 genes to verify that the above genes are BRCA-1 regulators according to methods disclosed above. Potential cleavage TST target sequences and the sequence of a ribozyme substrate recognition site for cleaving those TST target sequences is shown in tables 15-17 below.

Table 15: Validation RST and TST sequences for GenBank Acc. No. H07920

Validation ribozyme RST	TST target sequence
3'-AAGG UCAG ACAAAACG'-5' (SEQ ID NO:)	5'-TTCC AGTC TGTTTTGC-3' (SEQ ID NO:)
3'-GUAG CCAG UUCUCUUU-5' (SEQ ID NO:)	5'-CATC GGTC AAGAGAAA-3' (SEQ ID NO:)
3'- CUUG ACAG AUCUACCU-5' (SEQ ID NO:)	5'-GAAC TGTC TAGATGGA-3' (SEQ ID NO:)

Table 16: Validation RST and TST sequences for GenBank Acc. No. AA419229

Validation ribozyme RST	TST target sequence
5'-AGUG UCAG UACAGGGG-3' (SEQ ID NO:)	5'- TCAC AGTC ATGTCCCC -3' (SEQ ID NO:)
3'-GGGG UCAG AUUCAGGG-5' (SEQ ID NO:)	5'- CCCC AGTC TAAGTCCC -3' (SEQ ID NO:)
3'- UAAC UCAG AGCUCAGU-5' (SEQ ID NO:)	5'- ATTG AGTC TCGAGTCA -3' (SEQ ID NO:)

Table 17: Validation RST and TST sequences for GenBank Acc. No. Z49826

Validation ribozyme RST	TST target sequence
5'-UCGG UCAG ACUCUAUA-3' (SEQ ID NO:)	5'- AGCC AGTC TGAGATAT -3' (SEQ ID NO:)
3'-UGCC ACAG UUGACAGA-5' (SEQ ID NO:)	5'- ACGG TGTC AACTGTCT-3' (SEQ ID NO:)
3'- GGUC CCAG UUCGUGAC-5' (SEQ ID NO:)	5'- CCTG GGTC AAGCACTG -3' (SEQ ID NO:)
3'- ACCG UCAG UAGAGGUA-5' (SEQ ID NO:)	5'- TGGC AGTC ATCTCCAT -3' (SEQ ID NO:)
3'- GUAG UCAG UAAAGUGU-5' (SEQ ID NO:)	5'- CATC AGTC ATTTCACA -3' (SEQ ID NO:)

* * * * * * * * * * *

Throughout this application various publications have been referenced within parentheses. The disclosures of these publications in their entireties are hereby incorporated by reference in this application in order to more fully describe the state of the art to which this invention pertains.

Although the invention has been described with reference to the disclosed embodiments, those skilled in the art will readily appreciate that the specific experiments detailed are only illustrative of the invention. It should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

That which is claimed:

1. A substantially pure nucleic acid comprising a nucleotide sequence selected from SEQ ID NO:1 and a sequence of nucleotides having greater than 60% sequence identity to SEQ ID NO: 1, or a complementary sequence thereof.

2. The substantially pure nucleic acid of claim 1, wherein said nucleotide sequence comprises SEQ ID NO: 1.

3. A substantially pure nucleic acid comprising at least 16 nucleotides in length selected from SEQ ID NO:1 or selected from a sequence of nucleotides having greater than 60% sequence identity to SEQ ID NO: 1, or a complementary sequence thereof.

4. The substantially pure nucleic acid of claim 3, wherein said sequence is selected from SEQ ID NO:1.

5. A substantially pure nucleic acid comprising at least 18 nucleotides in length selected from SEQ ID NO:1 or selected from a sequence of nucleotides having greater than 60% sequence identity to SEQ ID NO: 1, or a complementary sequence thereof.

6. The substantially pure nucleic acid of claim 5, wherein said sequence is selected from SEQ ID NO:1.

7. A substantially pure nucleic acid comprising at least 20 nucleotides in length selected from SEQ ID NO:1 or selected from a sequence of nucleotides having greater than 60% sequence identity to SEQ ID NO: 1, or a complementary sequence thereof.

8. The substantially pure nucleic acid of claim 7, wherein said sequence is selected from SEQ ID NO:1.

9. A ribozyme comprising a target recognition sequence that selectively hybridizes to an RNA target selected from the group consisting of BR1 (SEQ ID NO: 1), ID4 (GenBank Acc. No. NM_001547), BBC1(GenBank Acc. No. X64707), RH3 (GenBank Acc. No. AL045940), and RH6 (GenBank Acc. No. AI276397), GenBank Acc. No. AA419229, GenBank Acc. No. H18950, GenBank Acc. No. H07920, GenBank Acc. No. H70047, GenBank Acc. No. H84815, GenBank Acc. No. AA757764, GenBank Acc. No. H19111, GenBank Acc. No. AA629897, GenBank Acc. No. AA663439, GenBank Acc. No. R15740, GenBank Acc. No. AA454570, GenBank Acc. No. AA485748, GenBank Acc. No. AA464601, GenBank Acc. No. AA102107: and GenBank Acc. No. Z49826, said hybridization resulting in cleavage of the RNA target.

10. The ribozyme of claim 9, wherein said target recognition sequence comprises the sequence N_8-AGAA- N_4.

11. The ribozyme of claim 10, wherein between 10-12 nucleotides of segments N_8 and N_5 taken together are complementary to the RNA target.

12. The ribozyme of claim 10, wherein said target recognition sequence is selected from the group consisting of:

5'–CCGGAUGCAGAACAAU–3', 5–AGUACAUUAGAAUACU–3'
5'–CUAGUGAGAGAAGGGA–3', 5'–UGAGAUCCAGAAAAGC–3'
5'–UGUUACUAGAAUGUU–3', and 5'–CCCUAUUUAGAAUUGU–3'

(SEQ ID NOS: 5–10, respectively).

13. A substantially pure nucleic acid target sequence tag recognized by a ribozyme, said sequence tag selected from the group consisting of:

5'–AUUGNGUCGCAUCCGG–3', 5'–AGUANGUCAAUGUACU–3',
5'–UCCCNGUCCUCACUAG–3', 5'–GCUUNGUCGGAUCUCA–3',
5'–AACANGUCAGUAACA–3, and 5'–ACAANGUCAAAUAGGG–3'

(SEQ ID NOS: 11–16, respectively), or a complementary sequence thereof.

14. A target recognition sequence of a ribozyme, said recognition sequence selected from the group consisting of:

5'–CCGGAUGCAGAACAAU–3', 5'–AGUACAUUAGAAUACU–3',
5'–CUAGUGAGAGAAGGGA–3', 5'–UGAGAUCCAGAAAAGC–3',
5'–UGUUACUAGAAUGUU–3', and 5'–CCCUAUUUAGAAUUGU–3'

(SEQ ID NOS: 5–10, respectively), or a complementary sequence thereof.

15. A substantially pure polypeptide that regulates the expression level of BRCA-1, comprising a sequence of amino acids shown in SEQ ID NO: 2 or a sequence of amino acids having greater than 50% sequence identity to SEQ ID NO: 2.

16. The substantially pure polypeptide of claim 15, wherein said sequence of amino acids is the sequence of amino acids shown in SEQ ID NO: 2.

17. A method for identifying a candidate gene for which the level of expression is affected by a BRCA-1 regulator, said method comprising:

a) hybridizing a first mRNA and a second mRNA to at least one candidate gene or portion thereof, wherein said first mRNA is obtained from cells expressing a ribozyme that targets and cleaves mRNA encoding a BRCA-1 regulator and wherein said second mRNA is obtained from control cells otherwise similar to those expressing the ribozyme except that the BRCA-1 regulator mRNA is not targeted by a ribozyme; and

b) comparing the relative amounts of the first and second mRNA that hybridizes to the gene, whereby a difference in the relative amounts of hybridization identifies a gene the level of expression of which is affected by a BRCA-1 regulator.

18. The method of claim 17, wherein said first and second mRNA is reverse transcribed into DNA before hybridization.

19. The method of claim 17, wherein said first and second mRNA or said gene is labeled with a detectable moiety.

20. The method of claim 19, wherein said detectable moiety is a fluorescent dye.

21. The method of claim 17, wherein two or more genes are hybridized to said first and second mRNA.

22. The method of claim 21, wherein said two or more genes are arrayed on a solid support before hybridization.

23. The method of claim 17, wherein said BRCA-1 regulator is selected from the group consisting of: BR1 (SEQ ID NO: 1), ID4 (GenBank Acc. No. NM_001547), BBC1(GenBank Acc. No. X64707), BR2 (GenBank Acc. No. AL045940), and BR3 (GenBank Acc. No. AI276397).

24. The method of claim 17, wherein said BRCA-1 regulator is selected from the group consisting of: GenBank Acc. No. AA419229, GenBank Acc. No. H18950, GenBank Acc. No. H07920, GenBank Acc. No. H70047, GenBank Acc. No. H84815, GenBank Acc. No. AA757764, GenBank Acc. No. H19111, GenBank Acc. No. AA629897, GenBank Acc. No. AA663439, GenBank Acc. No. R15740, GenBank Acc. No. AA454570, GenBank Acc. No. AA485748, GenBank Acc. No. AA464601, GenBank Acc. No. AA102107: and GenBank Acc. No. Z49826.

25. A method of identifying a compound that modulates the activity of a BRCA-1 regulator, said method comprising contacting a BRCA-1 regulator with a test compound and a target molecule responsive to the activity of the BRCA-1 regulator, wherein an increase or decrease in the activity of the BRCA-1 regulator for the target molecule in the presence of the test compound as compared to the absence of the test compound identifies a compound that modulates the activity of a BRCA-1 regulator.

26. The method of claim 25, wherein said test compound increases the activity of the BRCA-1 regulator.

27. The method of claim 25, wherein said test compound decreases the activity of the BRCA-1 regulator.

28. The method of claim 25, wherein said BRCA-1 regulator is selected from the group consisting of BR1 (SEQ ID NO: 1), ID4 (GenBank Acc. No. NM_001547), BBC1(GenBank Acc. No. X64707), BR2 (GenBank Acc. No. AL045940), and BR3 (GenBank Acc. No. AI276397).

29. The method of claim 25, wherein said BRCA-1 regulator is selected from the group consisting of: GenBank Acc. No. AA419229, GenBank Acc. No. H18950, GenBank Acc. No. H07920, GenBank Acc. No. H70047, GenBank Acc. No. H84815, GenBank Acc. No. AA757764, GenBank Acc. No. H19111, GenBank Acc. No. AA629897, GenBank Acc. No. AA663439, GenBank Acc. No. R15740, GenBank Acc. No. AA454570, GenBank Acc. No. AA485748, GenBank Acc. No. AA464601, GenBank Acc. No. AA102107: and GenBank Acc. No. Z49826.

30. The method of claim 25, wherein said BRCA-1 regulator activity is DNA binding activity.

31. The method of claim 25, wherein said activity measured is expression of the nucleic acid element or a gene through its operative linkage to a reporter gene.

32. The method of claim 25, wherein said BRCA-1 regulator activity is protein kinase activity.

33. The method of claim 25, wherein said BRCA-1 regulator activity is GTP binding activity.

34. The method of claim 25, wherein said BRCA-1 regulator activity is protease activity.

35. The method of claim 25, wherein said BRCA-1 regulator activity is protein binding activity.

36. The method of claim 25, wherein said BRCA-1 regulator activity is hormone binding activity.

37. A method of treating cancer, comprising contacting a cancer cell with an expression vector that encodes a ribozyme selectively reactive with an RNA

encoding a BRCA-1 regulator or contacting the cancer cell with the ribozyme, wherein said BRCA-1 regulator is selected from the group consisting of: BR1 (SEQ ID NO: 1), ID4 (GenBank Acc. No. NM_001547), BBC1(GenBank Acc. No. X64707), BR2 (GenBank Acc. No. AL045940), and BR3 (GenBank Acc. No. AI276397).

38. The method of claim 37, wherein said BRCA-1 regulator is selected from the group consisting of: GenBank Acc. No. AA419229, GenBank Acc. No. H18950, GenBank Acc. No. H07920, GenBank Acc. No. H70047, GenBank Acc. No. H84815, GenBank Acc. No. AA757764, GenBank Acc. No. H19111, GenBank Acc. No. AA629897, GenBank Acc. No. AA663439, GenBank Acc. No. R15740, GenBank Acc. No. AA454570, GenBank Acc. No. AA485748, GenBank Acc. No. AA464601, GenBank Acc. No. AA102107: and GenBank Acc. No. Z49826.

39. The method of claim 37, wherein said ribozyme comprises a target recognition sequence selected from the group consisting of

5'–CCGGAUGCAGAACAAU–3', 5'–AGUACAUUAGAAUACU–3', 5'–CUAGUGAGAGAAGGGA–3', 5'–UGAGAUCCAGAAAAGC–3', 5'–UGUUACUAGAAUGUU–3', and 5'–CCCUAUUUAGAAUUGU–3' (SEQ ID NOS: 5-10, respectively).

40. A method of detecting a neoplastic cell in a sample wherein said neoplastic cell is associated with an altered expression of a BRCA-1 regulator or an altered structure of a BRCA-1 regulator as compared to a normal cell, comprising:

(a) contacting the sample with a detectable agent specific for a BRCA-1 regulator nucleic acid or an encoded BRCA-1 regulator polypeptide; and

(b) detecting said nucleic acid or polypeptide in said sample, wherein altered expression or structure of said polypeptide indicates the presence of a neoplastic cell in said sample.

41. The method of claim 40, wherein said BRCA-1 regulator is selected from the group consisting of BR1 (SEQ ID NO: 1), ID4 (GenBank Acc. No. NM_001547), BBC1(GenBank Acc. No. X64707), BR2 (GenBank Acc. No. AL045940), and BR3 (GenBank Acc. No. AI276397).

42. The method of claim 41, wherein said BRCA-1 regulator is selected from the group consisting of: GenBank Acc. No. AA419229, GenBank Acc. No. H18950, GenBank Acc. No. H07920, GenBank Acc. No. H70047, GenBank Acc. No. H84815, GenBank Acc. No. AA757764, GenBank Acc. No. H19111, GenBank Acc. No. AA629897, GenBank Acc. No. AA663439, GenBank Acc. No. R15740, GenBank Acc. No. AA454570, GenBank Acc. No. AA485748, GenBank Acc. No. AA464601, GenBank Acc. No. AA102107: and GenBank Acc. No. Z49826.

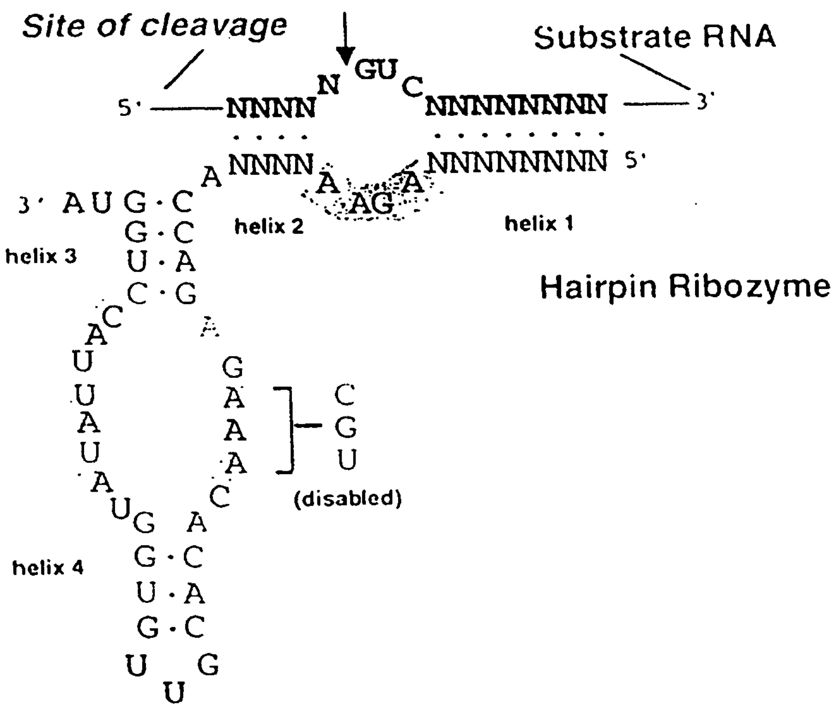
Site of cleavage
Substrate RNA
5'
NNNN
N
GU
C
NNNNNNNNN
3'
NNNN
NNNNNNNNN
5'
A
A
AG
A
3' AUG · C
G · C
U · A
C · G
helix 2
helix 1
helix 3
Hairpin Ribozyme
C
A
U
U
A
U
A
U
G
A
G
A
A
A
C
C
G
U
(disabled)
helix 4
G · C
U · A
G · C
U
U
G

FIGURE 1

```
BR1 348  at-atatgctgctataaaataa-----g---agcaagtttaaaaagatta--aag-c-ca 394
    226  atcata-gcagct-tacaacaacacgggtacagcca--tcaagaaga--aggaagacaca 279
    76    I  I   A  A   Y  N  N  T  G  T  A    I  K  K    K  E  D  T

    395  cat-g--tgccgcttgggaag--aa-attgcagtt-ctc--taccaggct--taatagga 443
    280  catagaatg--gc---ggaagccaacagggcatttgctcattac-agg-tggtgataagc 332
    94    H  R  M   A     E  A  N  R  A  F  A  H  Y   R   W ^^^

    444  aatgatatca-aatcccttcattccatcatcaat-     -cctcccata-gctaaaatccg 496
    333  a-tg----cagaatg--tggaagcgag-at--atgttgacctgg-ataagcttaga---- 377
    1     M      Q  N   V  E  A  R    Y  V  D  L   D  K  L  R

    497  taatattggaattatggctcatattgatgcaggcaaaactaccaccacagaaagaatatt 556
    378  -aatatcggtataatggctcatatcgacgccggaaaaactactacaacagagagaatact 436
    16    N  I  G  I  M  A  H  I  D  A  G  K  T  T  T  T  E  R  I  L

    557  gtactatt---ccggatatacaaga--tcactgggagatgttgatgatggagaca----c 607
    437  ---ctattacacggga-agaaaacacttca-taggagacgtcgatgaaggaaacacaacc 491
    36       Y  Y  T  G   R  K  H  F   I  G  D  V  D  E  G  N  T  T

    608  agtgac-agat--ttcatggcccaagagcgagaaagaggcattactattcaatcag-ctg 663
    492  actgactggatgcctca-gg---aaaag---gaaagaggcattacgat--a--cagtctg 540
    54    T  D  W  M  P  Q      E  K     E  R  G  I  T  I       Q  S

    664  ctgttaca----tttgattggaaaggttatagagtcaatctaattgatacaccaggtca 718
    541  ctgccacaacgtgttt--ctggaaaggttaccgaatcaacataattgatacacccggaca 598
    70   A  A  T  T  C  F    W  K  G  Y  R  I  N  I  I  D  T  P  G  H

    719  tgtggactttaccttg---gaggttg--agcggtgcctaagagtgt--tggatggtgcag 771
    599  cgttgactttac---ggccgaggttgaaag-ggcgctt---cgtgttctcgatggagcga 651
    90    V  D  F  T     A  E  V  E  R  A  L    R  V  L  D  G  A

    772  t--ggctgtatttgatgcctctgctggtgtagaggcccagactctc----acagtatgga 825
    652  tacg--tgttttcgacgccacggcaggagtcgagcctcag----tcggaaaccgtctgga 705
    107  I  R    V  F  D  A  T  A  G  V  E  P  Q     S  E  T  V  W

    826  ggcaagctgataaacacaatatacctcga--atctgtttttaaacaagatggacaaaac 883
    706  gacaggctgacaagtacaatgt-cc-cgaggatcgctttcatgaacaagatggacaaagt 763
    125  R  Q  A  D  K  Y  N  V    P  R  I  A  F  M  N  K  M  D  K  V

    884  tggagcaagctt--taagtat-gcagttgaaagcatc----agagagaagttaaaggcaa 936
    764  gggagc-agatttcta--catggccgttgaaa-c-tcttgta-aca-aaactcagggcaa 816
    145   G  A   D  F  Y    M  A  V  E   T  L  V   T   K  L  R  A
```

FIGURE 2A

```
937  agcctttgcttt----tacagttaccaattg---gtgaagccaaa-actttcaaaggagt 988
     | ||     | ||     | ||| | || || |     ||| ||   ||| || ||| | || ||
817  atcc----cattccggtgcagatgccgatcggcagtg-ag--aaagac-ttccagggtgt 868
162  N  P       I  P  V  Q  M  P  I  G  S  E    K  D  F  Q  G  V

989  ggtgga--t-gtagtaatgaaagaaaaacttct-ttggaattgcaattca---aatgatg

     | | ||  |  ||  |||||||| |     ||| |||     |     |||   || ||||
869  gattgacctcataaaaatgaaagcga----tctactgg----g----tcagcgaa-gatg 915
180   I  D  L  I  K  M  K  A      I  Y  W          V  S  E  D

1042 gaaaagacttttg-a-ga-gaaagcccctctt-----ggaaatgaatgatcctgaattgc

     ||     || || | || |||||    | ||     |||| | |  ||    |||
916  ga----tctgtgtacgaggaaag-agatattccggaggaactca--ga---gaa-----  959
195  G      S  V  Y  E  E  R  D  I  P  E  E  L    R     E

1094 gaag--gaaacaactgaagcaag--ga-atgccttaattgaacaagttgcagatttggat

     ||||  ||||  |  ||||  ||  || ||||      ||||| || | |||||  | ||
960  gaagcagaaatga--gaag--agaagagatgc-----ttgaaaaatcgcagaactcgac

210   E  A  E  M    R  R    E  E  M       L  E  K  I  A  E  L  D

1145 gatg-aat--ttg--------ct-ga---cttgg---ttt---tagaagaatttagtgag

     || | ||| |||         ||| ||   |  ||   |||    ||||||| ||  |||
1011 gaagaaattcttgaaaaatacctcgaaggcgcggaaatttccgaagaagaa-ataaagag

227   E  E  I  L  E  K  Y  L  E  G  A  E  I  S  E  E  E   I  K  R

1188 aactttgatttgttaccagctgaaaagctacagactgcaatacatagagtgacactagct

      | ||||             ||||||| ||            ||||||   | ||
1070 ggtcttga------------gaaaagccac-----------cataga---aaac-----

247   V  L               R  K  A  T             I  E     N

1248 cagacagcagtgcctgtgctttgtggaagtgccctgaaa----aacaaagggatacagcc

        | |||||| || || || ||||| || | |   |||     ||||| || ||||||||
1098 ---aaagcagttcccgttctctgtgg-agcggc---aaaagccaacaagggaatacagcc

256     K  A  V  P  V  L  C  G  A  A    K  A  N  K  G  I  Q  P

1304 -cttgttagatgc--tgtt--actat-----gt----actt-acct-tcacctgaagagc

      |||  | || ||  || |  |||||      ||     ||| | ||  |||| | | ||
1151 gctt-ctggacgcggtgatagactatctgccgtcgcctcttgatctaccaccggtaaagg

273   L  L  D  A  V  I  D  Y  L  P  S  P  L  D  L  P  P  V  K

1348 gtaactatgaa---tttctg-c--agtggt-------ataa----ggatga---------

     |     ||| |   |||| | |  || |||        | ||    ||||||
1210 g-----atggagggtttccgacggagaggttgtctacagaaagccggatgaaaacgaacc
```

FIGURE 2A continued

```
293  G      W  R  V  S  D  G  E  V  V  Y  R  K  P  D  E  N  E  P

1382 ctt---atgtgcattggcatttaaagttc-------tccatgacaagcagcgaggaccac
     |||   |   ||  |||  |||||||| |        |||||| |||       |||| ||
1265 cttcaca---gccctggtctttaaagtgcaggtggatccat-aca------taggaaaac

312    F  T     A  L  V  F  K  V  Q  V  D  P  Y        I  G  K

1432 tggttt--ttatgcgcatttactcagg--cact--ataaa--accccagttggccattca
     | || |  ||  | |  | ||||| ||   |||  | |||  |     ||||   | || |
1315 tcgtgtacttcagag--tctactctggaagactggagaaagga----agtt--acgttta

328  L  V  Y  F  R     V  Y  S  G  R  L  E  K  G        S     Y  V  Y

1484 taa-t--attaatggaaactgcacggagagaata---ag--tcgtctgcttttgc-cgtt
     || |  |  || ||  | | | ||||||||||  ||  ||||||  |  ||| |
1367 caactccacaaaggg--tcag-agggagagaatatcgaggatcgtct--tcatgcac---

346    N  S  T  K  G     Q  R  E  R  I  S  R  I  V     F  M  H

1535 tgctgaccaacatgtaga--aatcccttcattgactgctggtaacattgctttga-ctgt
     || ||    | | |||  ||    || ||| ||  |     ||       || | |
1419 -gcaga-caaga-g-agaggaa---gttgatt-act--t-----ca-------gacccgg

363    A  D  K  R  E  E     V  D  Y  F                 R  P  G

1592 tgggcttaaacatactgc -cactggagacaccattgtctcatccaagtccagtgcatta
     ||       | |||  ||  |  ||||  |     ||||   ||||    ||   ||
1457 tg-------atata--gcggca--ggagtc-----ggtct----caag----gt---tt-

376              D  I     A  A     G  V     G  L     K     V

1650 gctgcag-ctcgtagagccgaacgggagggagaaaagaagcacagacaaaacaat-gaag
     || ||| || | |||  ||  | ||   ||| |||||   ||  | ||  |  | |||
1489 -ct-cagact-ggagacacgctctggcacgag-aaaga---acctataatcc--tcgaa-

386    S  Q  T  G  D  T  L  W  H  E  K  E     P  I  I     L  E

1708 cagaga--gacttttattggctggagtggagattccagaacctgtttct-tctgtacca
     | |||  |||||        |              |||||||| ||| ||| ||| | ||
1539 -a-agatcgactt-------c-------------ccagaacccgttatctctct-tgccg

403    K  I  D  F                         P  E  P  V  I  S  L  A

1765 tagaaccccca-tcactgtctaagcagccagatttg--gaacatgcgt--tgaa---a-t
     |||||   ||| |||| |  ||  || |||| ||  ||| | || |  ||||   | |
1576 tagaa---ccagtcac-g---aa--ag-caga--tgaagaa-aagc-tcgtgaaggcact

415  V  E     P  V  T     K     A  D     E  E  K  L  V  K  A  L
```

FIGURE 2A continued

```
1816 g-----tc-ttcagcgtgaagatcc--cagtttgaaagtgaggctagatcctgactctgg
     |        || ||| |    ||||| ||  || ||   | |||||| |        ||||
1622 gctggctctttctg-aagaagaccccacactt--caggtgagggt------tgac-----
430   L  A  L  S   E  E  D  P  T  L   Q  V  R  V        D

1868 acaagcaccac-ccagaagctgtccacatcaatt-----at-----a--ttcatcttcca
     | ||| |  ||  ||||      |||||| ||    ||    |  ||||||||
1668 a-aagaa--acgggagaa------aacatcatttcaggaatgggcgagcttcatctt---
446   K  E   T  G  E        N  I  I  S  G  M  G  E  L  H  L

1915 ttggccagaaatttggctttg-c ggttg-gga-agtgtt-ggtatttagctaagtgttg
            |||| | | | | | | ||||| || || ||| ||| |  |  | || ||||
1716 -------gaaa-tagtcgtagacaggttgaagagagagttcggtgtcaacgtgagagttg
462         E   I  V  V  D  R  L  K  R  E  F  G  V  N  V  R  V

1970 gcatcattagccttagtcagatgctacttgcccagcagataattcaactatcaaagaa--
     | | |   ||||    |||| ||   |||   ||| ||      ||||||| |||||
1768 g-a-c---agcc----tcaggtg--gctt--acag-gga------aactatc-aagaaat
479  G     Q  P    Q  V    A   Y  R  E        T  I   K  K

2028 -----gaag--ga--gagaat------ggc---caggt-----aattcattgtttattga
           |||  .  :           .  .  :  .|    |  :.  :
180[illegible] ctgctgaagccgaaggaaaatacatcaggcagaccggtggtagaggtca -g tacgg-
492  S  A  E  A  E  G  K  Y  I  R  Q  T  G  G  R  G  Q     Y  G

2065 ttcat-t-attccagtcagaaaag 2086
      |||| | ||||   |||||| ||
1862 -tcatgtgattc---tcagaatag 1881
511    H  V  I    L  R  I
```

FIGURE 2A continued

```
BR1  MAQERERGITIQSAAVTFDWKGYRVNLIDTPGHVDFTLEVERCLRVLDGAVAVFDASAGV
     M QE+ERGITIQSAA T   WKGYR+N+IDTPGHVDFT EVER LRVLDGA+ VFDA+AGV
 57  MPQEKERGITIQSAATTCFWKGYRINIIDTPGHVDFTAEVERALRVLDGAIRVFDATAGV 116

BR1  EAQTLTVWRQADKHNIPRICFLNKMDKTGASFKYAVESIREXXXXXXXXXXXXXIGEAKTF
     E Q+ TVWRQADK+N+PRI F+NKMDK GA F  AVE++                IG  K F
117  EPQSETVWRQADKYNVPRIAFMNKMDKVGADFYMAVETLVTKLRANPIPVQMPIGSEKDF 176

BR1  KGVVDVVMKEKLLWNCNSNDGKDFERKPLLEMNDPELLKETTEARNALIEKVADLDLEEA
     +GV+D++  + + W   S DG  +E + + E    EL +E   R  ++E++A+LD+E
177  QGVIDLIKMKAIYW--VSEDGSVYEERDIPE----ELREEAEMRREEMLEKIAELDEE-- 226

BR1  DLVLEEFSENFDLLPAEKLQTAIHRVTLAQTAVPVLCGSALKNKGIQPLLDAVTMYLPSP
       +LE++ E  + +  E+++  + + T+  AVPVLCG+A NKGIQPLLDAV  YLPSP
229  --ILEKYLEGAE-ISEEEIKRVLRKATIENKAVPVLCGAAKANKGIQPLLDAVIDYLPSP 285

BR1  EE---------RNYEFLQWYKDD---LCALAFKVLHDKQRGPLVFMRIHSGTIKPQLAIH
      +          + E +    D+     AL FKV  D   G LV+ R+-SG ++     ++
286  LDLPPVKGWRVSDGEVVYRKPDENEPFTALVFKVQVDPYIGKLVYFRVYSGRLEKGSYVY 345

BR1  HINGNCTERISPLLLFFADQHVEIPSLTAGNIALTVGLKHTATGDTTVXXXXXXXXXXX
     N       ERISR++    AD+  E+      G+IA  VGLK + TGDT
     NSTKGQRERISFIVFKHADKREEVDYFRPGDIAAGVGLKVSQTGDTLX------------ 393

BR1  XXXXXXXXXHRLDEAF+LLLAGVEIPEPVLFCTIEIISLSKQPDIEHALSIQREDISLK
                   +E F +IL  ++ PEPV+    +EP + + I  L  AL  L  EDP+L+
394  -------------HEKEPIILEKIDFPEPVISLAVEPVTKADEEKLVKALLALSEEDPTLG 441

BR1  VRLDPDSGQ
     VR+D ++G+
442  VRVDKETGE 450
```

FIGURE 2b

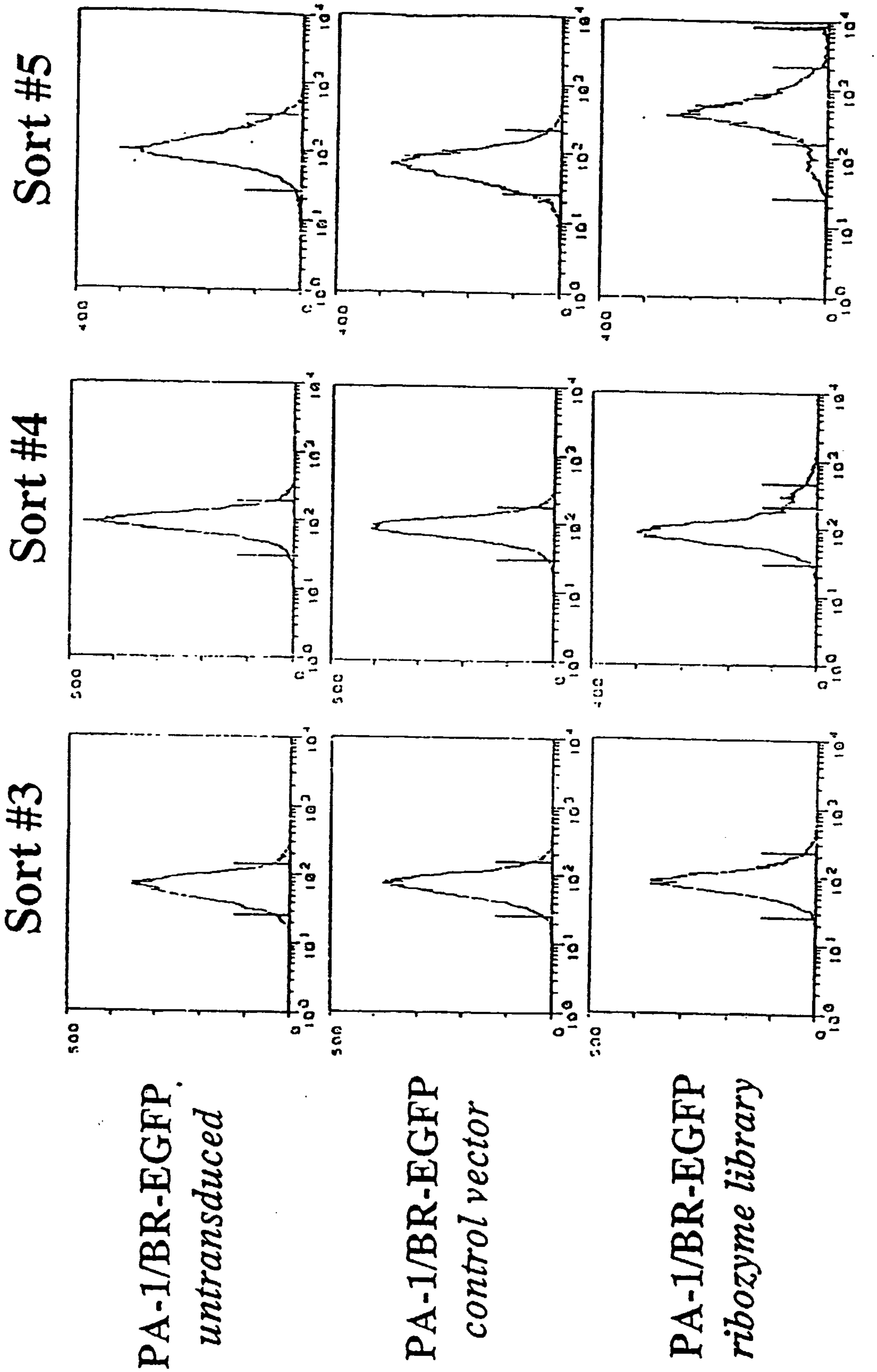
Sort #3
Sort #4
Sort #5
PA-1/BR-EGFP
untransduced
PA-1/BR-EGFP
control vector
PA-1/BR-EGFP
ribozyme library

FIGURE 3

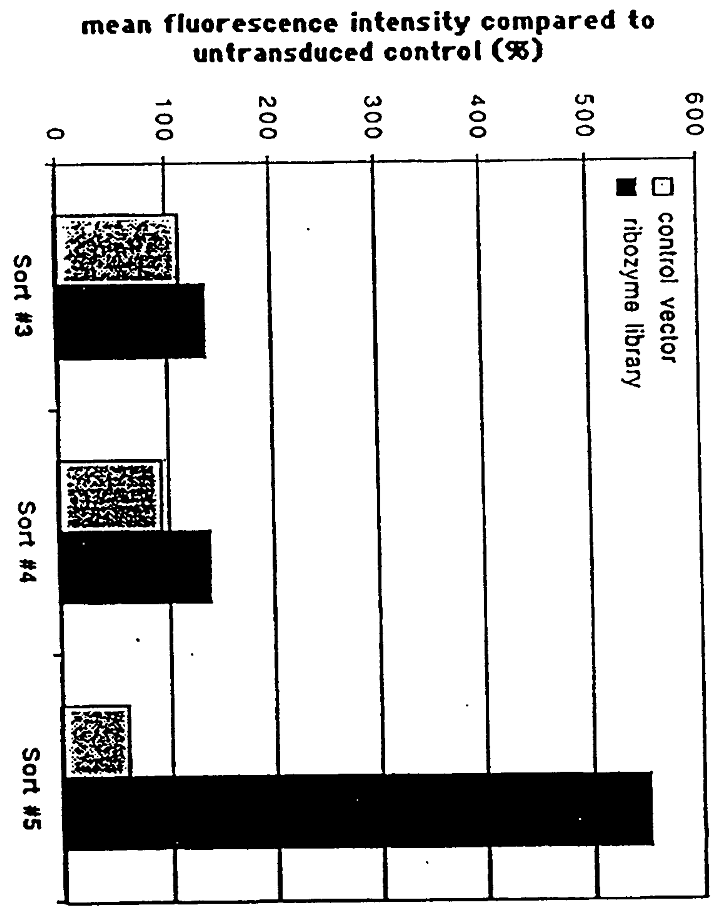
mean fluorescence intensity compared to untransduced control (%)
0
100
200
300
400
500
600
control vector
ribozyme library
Sort #3
Sort #4
Sort #5

FIGURE 4

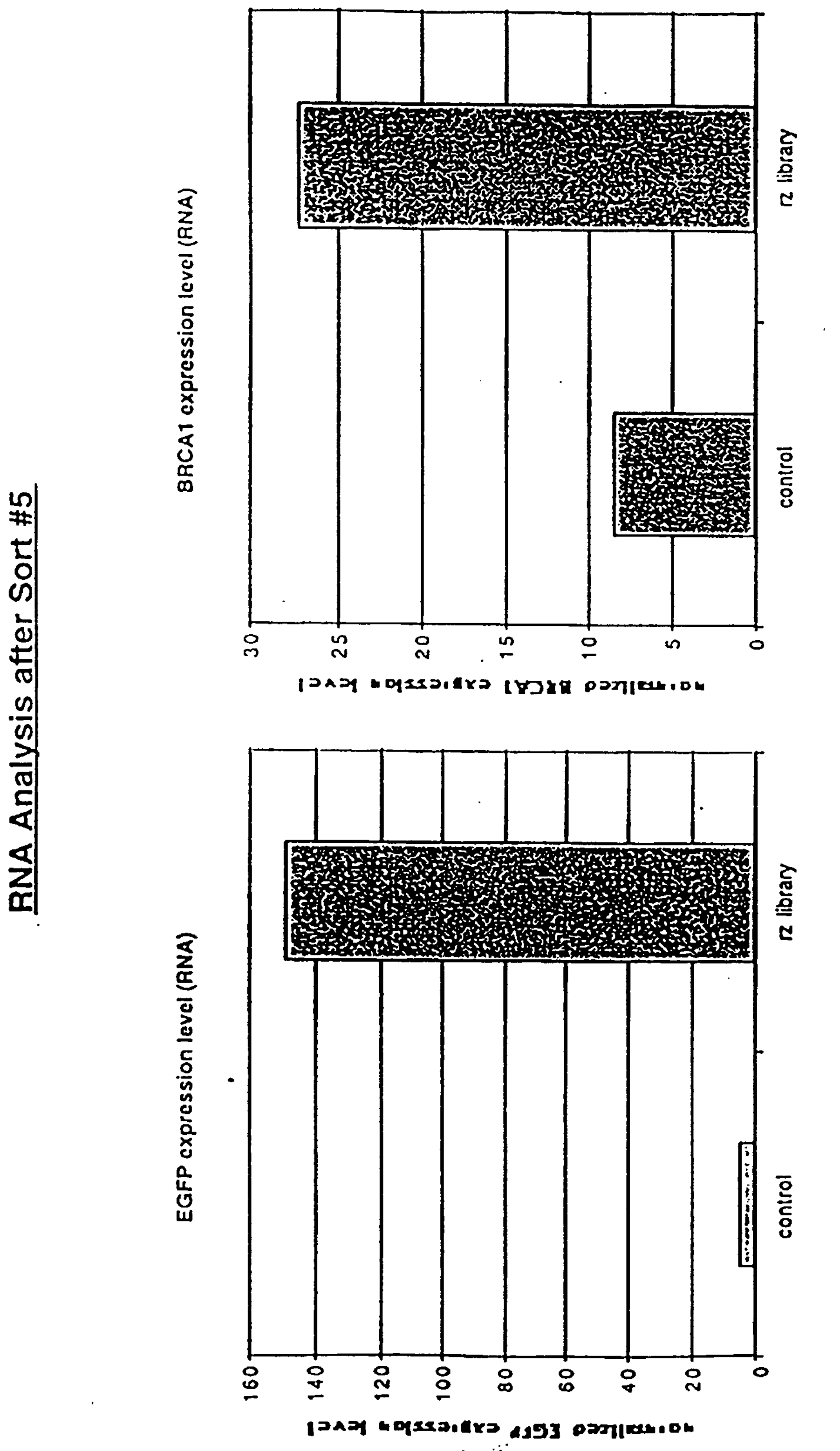
RNA Analysis after Sort #5
BRCA1 expression level (RNA)
0
5
10
15
20
25
30
control
rz library
EGFP expression level (RNA)
0
20
40
60
80
100
120
140
160
control
rz library

FIGURE 5

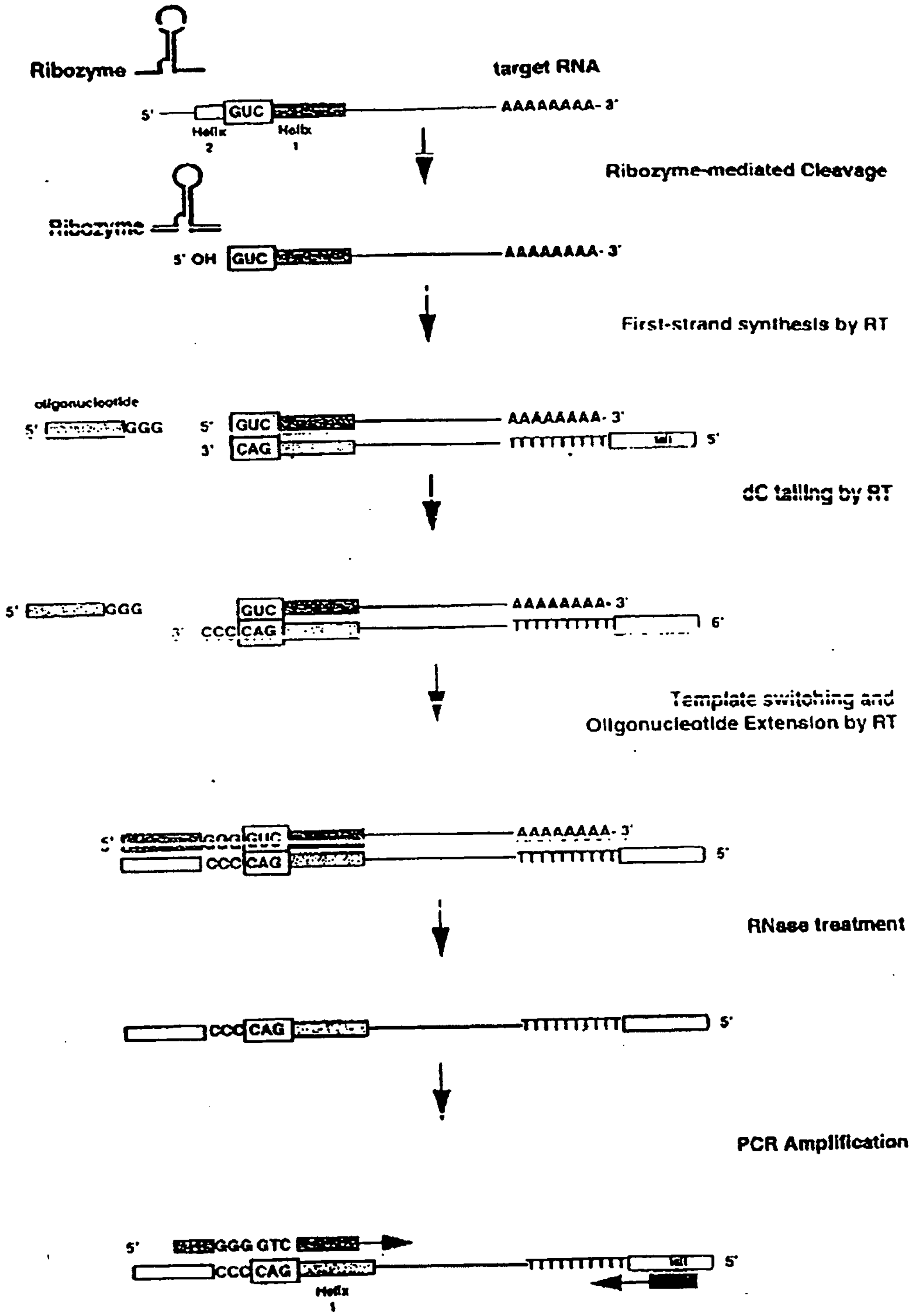
Ribozyme
target RNA
5'
GUC
AAAAAAAA- 3'
Helix 2
Helix 1
Ribozyme-mediated Cleavage
Ribozyme
5' OH
GUC
AAAAAAAA- 3'
First-strand synthesis by RT
oligonucleotide
5'
GGG
5'
GUC
AAAAAAAA- 3'
3'
CAG
5'
dC tailing by RT
5'
GGG
GUC
AAAAAAAA- 3'
CCC
CAG
Template switching and
Oligonucleotide Extension by RT
5'
GUC
AAAAAAAA- 3'
CCC
CAG
5'
RNase treatment
CCC
CAG
5'
PCR Amplification
5'
GGG GTC
CCC
CAG
5'
Helix 1

FIGURE 6

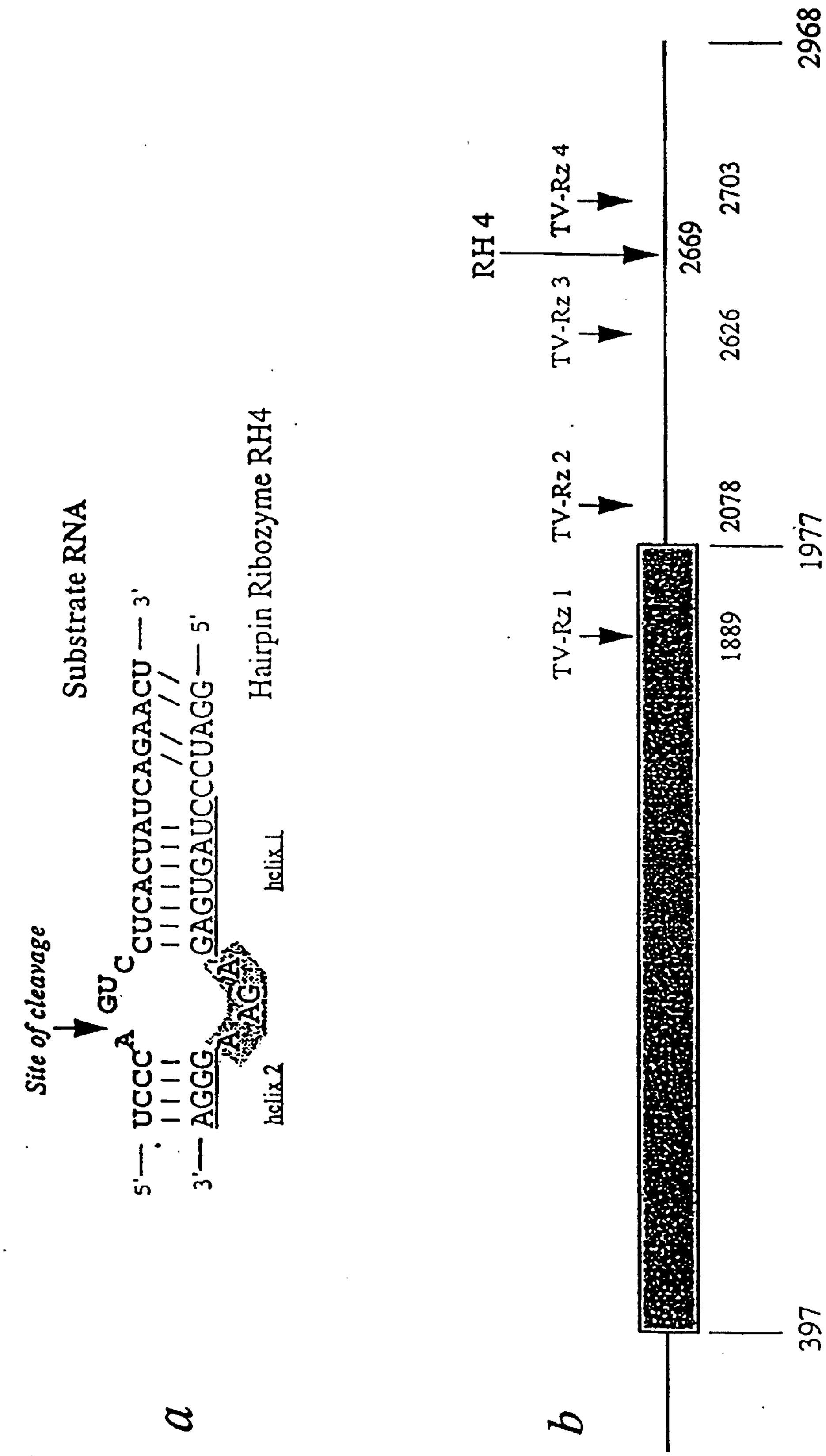
a
Site of cleavage
Substrate RNA
5'— UCCC A GU C CUCACUAUCAGAACU — 3'
3'— AGGG GAGUGAUCCCUAGG — 5'
helix 2
helix 1
Hairpin Ribozyme RH4
b
TV-Rz 1
TV-Rz 2
TV-Rz 3
RH 4
TV-Rz 4
397
1889
1977
2078
2626
2669
2703
2968

FIGURE 7

BBC1

TV6 gcatgGTCttgaagcc
TV7 cccgcGTCgggtcccg
TV8 gtcggGTCccatccgg
aacaaGTCcacggagt
acggaGTCcctgcaga
gaccgGTCatgcccgt
TV9 tgcccGTCcggaacgt
ggaacGTCtataagaa
ctcgaGTCatcactga
cgctaGTCtccgtatg
aatcaGTCgggcagtc
TV10 gggcaGTCatgctggg
gctggGTCtccacgtg
agacaGTCcaagccct

CHLR2

agctgGTCctcgccgaat
ctgtaGTCggacacac
TV11 ggctgGTC tctctcgg
ggcctGTCcct attac
gtgctGTCctatcaga
tggagGTCagcg gctc
acagaGTCtgtcacag
cgaga GTCaggccag c
ccag cGTCccgcaa cc
cctcaGTC agagcacc
c agtgGTCatt gccgg
cc ggtGTCtgacttcc
c tgtgGTCacg tgatc
TV12 t cgggGTCtccaacca
TV13 ggtggGTCgcatcc tc
TV14 gggtgGTCtg tttctt
gcca gGTCcatgcc ca
atcagGTC tgcagagc
aggccGTCa accagtc
a accaGTCcat aggca
TV15 cccctGTCctgg ccaa

ID4

gctta GTCggagctc c
TV19 tg agtGTCgccc
TV20 ctc ccGTCaattg ttg

FIGURE 8

ggagt**GTC**gcg gtgcc
TV21 gcgct**GTC**caggtgtg
agaaa**GTC**a gcaaagt
TV22 gacct**GTC**c agccgcg
ctttt**GTC** tcttcatt

AF6

TV23 ttcgg**GTC**tcta gtac
atgct**GTC** ctctccca
tgca t**GTC**agcgga
ctgct **GTC**tactaca g
gt atg**GTC**tgga aaaa
tgagt**GTC** ctttacaa
t ttta**GTC**tttcagtt
gac gg**GTC**tgtct atg
ggtct**GTC**t atggctc
gagtg**GTC**a ctgtgac
ggggc **GTC**ccatgtat
agag t**GTC**gtctgc ct
ct aca**GTC**cact tt
aagtt**GTC** ccctacat
ttgca**GTC** gtcaacaa
agggt**GTC**atccaga a
TV26 cctta**GTC** ggatcaca
gcca t**GTC**cttgct ac
gcaga**GTC**aatg ccgc
tatcc **GTC**agcagtt g
gg act**GTC**atct gagc
a tgtt**GTC**aga aacat
t gaaa**GTC**acc ttctg
tgctg**GTC** aagataaa
gtgaa**GTC**gg ttgtga
tg gac**GTC**tagc tgca
acgaa**GTC**tg gttgga
TV27 ccacg**GTC**tgg ccacc
tccta**GTC**c tggaggg
tct aa**GTC**ggtca ccc
agtcg**GTC**a cccctgc
tggcc**GTC**tcccagcc
ggaat**GTC** catggatt
ctgc c**GTC**tgcgca gg
act ca**GTC**cttaa acc
aaaca**GTC**attcggg a

FIGURE 8 continued

ggaca **GTC**tgtcccc c
a gtct**GTC**ccc cgacc
cagga**GTC** gaagcaga
ctgtg**GTC**cctattt t

BR1

acaaa**GTC**tgcaacgg
ctcaa**GTC**tcgacacc
aatga**GTC**atcagaca
ataga**GTC**aatctaat
accag**GTC**atgtggac
aataa**GTC**gtctgctt
ccatt**GTC**tcatccaa
tccaa**GTC**cagtgcat
tcact**GTC**taagcagc
gaaat**GTC**ttcagcgt
TV1 aagct**GTC**cacatcaa
cctta**GTC**agatgcta
TV2 ttcca**GTC**a gaaaagtc
gaaaa**GTC**aaaagcaa
tttct**GTC**tttctttc
tcata**GTC**actaatgt
caagt**GTC**cttttcag
TV3 cagtg**GTC**atggaata
ctgaa**GTC**gatggctt
TV4 ggctt**GTC**ccagtcct
tccca**GTC**ctcactat
aacta**GTC**taagaatt
TV5 ttgaa**GTC**aatgtgtt
agaaa**GTC**agtttcag

BR3

TV 17gaac tGTCaagaaata
TV 16aattt GTCttaaata g
aacaaGTCagtaacaa

BR4

agaa a**GTC**caaaaatc
TV 33 ccgt g**GTC**tggaagga
TV34 ctcaa**GTC**agggccaa

FIGURE 8 continued

agca cGTCtcccac gt
agtaaGTCa atgtacc
ccgaa GTCaccagatg
TV 35 gggca GTC tagaaaaa
accga GTCctgcttg c
TV 36 cgcacGTCc gagacaa
TV37 tgat gGTCttgacg at

FIGURE 8 continued

1 TCTCACCGTGGGCTCTTGAAGCAGGAGGCTCGACAAAGTCTGCAACGGCCTCAAGTCTCGACACCGCCGGAAGTGTGGAG 80

81 GTGTGGCTGCCCGAGGCGGTAGAGCGGAGAAAGCCTATTTTCTTTTCGTTTAGATACATTGCCTTTTGCCTAGGCTGGCG 160

161 TCGAGACTTGAGGCCGTTGCAGACTTTGGCGCGGCTCGCGCCTCCTGCTTCAAGAGCCCAGCGGTGAGAGCTGGCCTGCG 240

241 GCACGCGGCCTA ATG CCA GAC AGT AAC AGT TTG GAG GAT CAA GAT GTT GAC CAA CTT GAG GAT 303
1 M P D S N S L E D Q D V D Q L E D 17

304 ATT TGC AAT GAG TCA TCA GAC AAT ACC CAG TGT GTA TAT TAA TAATAT ATG CTG CTA TAA A 364
18 I C N E S S D N T Q C V Y * M L L * 4

365 ATAAGAGCAAGTTTAAAAAGATTAAAGCCAC ATG TGC CGC TTG GGA AGA AAT TGC AGT TCT CTA CCA 431
1 M C R L G R N C S S L P 12

432 [illegible] TTA ATA GGA AAT GAT ATC AAA [illegible] [illegible] CAT TCA ATC ATC AAT [illegible] [illegible] ATA [illegible] AAA 491
13 G L I G N D I K S L H S I I N P I I A E 32

492 ATC CGT AAT ATT GGA ATT ATG GCT CAT ATT GAT GCA GGC AAA ACT ACC ACC ACA GAA AGA 551
33 I R N I G I M A H I D A G K T T T T E R 52

552 ATA TTG TAC TAT TCC GGA TAT ACA AGA TCA CTG GGA GAT GTT GAT GAT GGA GAC ACA GTG 611
53 I L Y Y S G Y T R S L G D V D D G D T V 72

612 ACA GAT TTC ATG GCC CAA GAG CGA GAA AGA GGC ATT ACT ATT CAA TCA GCT GCT GTT ACA 671
73 T D F M A Q E R E R G I T I Q S A A V T 92

672 TTT GAT TGG AAA GGT TAT AGA GTC AAT CTA ATT GAT ACA CCA GGT CAT GTG GAC TTT ACC 731
93 F D W K G Y R V N L I D T P G H V D F T 112

732 TTG GAG GTT GAG CGG TGC CTA AGA GTG TTG GAT GGT GCA GTG GCT GTA TTT GAT GCC TCT 791
113 L E V E R C L R V L D G A V A V F D A S 132

792 GCT GGT GTA GAG GCC CAG ACT CTC ACA GTA TGG AGG CAA GCT GAT AAA CAC AAT ATA CCT 851
133 A G V E A Q T L T V W R Q A D K H N I P 152

852 [illegible] ATC TGT TTT TTA AAC AAG ATG GAT AAA ACT GGA GCA AGC TTT AAG TAT GCA GTT GAA 911
153 P I C F L N K M D K T G A S F F Y A V E 172

912 AGC ATC AGA GAG AAG TTA AAG GCA AAG CCT TTG CTT TTA CAG TTA CCA ATT GGT GAA GCC 971
173 S I R E K L K A K P L L L Q L P I G E A 192

972 AAA ACT TTC AAA GGA GTG GTG GAT GTA GTA ATG AAA GAA AAA CTT CTT TGG AAT TGC AAT 1031
193 K T F K G V V D V V M K E K L L W N C N 212

1032 TCA AAT GAT GGA AAA GAC TTT GAG AGA AAG CCC CTC TTG GAA ATG AAT GAT CCT GAA TTG 1091
213 S N D G K D F E R K P L L E M N D P E L 232

1092 CTG AAG GAA ACA ACT GAA GCA AGG AAT GCC TTA ATT GAA CAA GTT GCA GAT TTG GAT GAT 1151
233 L K E T T E A R N A L I E Q V A D L D D 252

1152 GAA TTT GCT GAC TTG GTT TTA GAA GAA TTT AGT GAG AAT TTT GAT TTG TTA CCA GCT GAA 1211
253 E F A D L V L E E F S E N F D L L P A E 272

1212 AAG CTA CAG ACT GCA ATA CAT AGA GTG ACA CTA GCT CAG ACA GCA GTG CCT GTG CTT TGT 1271
273 K L Q T A I H R V T L A Q T A V P V L C 292

1272 GGA AGT GCC CTG AAA AAC AAA GGG ATA CAG CCC TTG TTA GAT GCT GTT ACT ATG TAC TTA 1331
293 G S A L K N K G I Q P L L D A V T M Y L 312

1332 CCT TCA CCT GAA GAG CGT AAC TAT GAA TTT CTG CAG TGG TAT AAG GAT GAC TTA TGT GCA 1391
313 P S P E E R N Y E F L Q W Y K D D L C A 332

1392 TTG GCA TTT AAA GTT CTC CAT GAC AAG CAG CGA GGA CCA CTG GTT TTT ATG CGC ATT TAC 1451
333 L A F K V L H D K Q R G P L V F M R I Y 352

1452 TCA GGC ACT ATA AAA CCC CAG TTG GCC ATT CAT AAT ATT AAT GGA AAC TGC ACG GAG AGA 1511
353 S G T I K P Q L A I H N I N G N C T E R 372

1512 ATA AGT CGT CTG CTT TTG CCG TTT GCT GAC CAA CAT GTA GAA ATC CCT TCA TTG ACT GCT 1571

FIGURE 9

373 I S R L L P F A D Q H V E I P S L T A 392

1572 GGT AAC ATT GCT TTG ACT GTT GGG CTT AAA CAT ACT GCC ACT GGA GAC ACC ATT GTC TCA 1631
393 G N I A L T V G L K H T A T G D T I V S 412

1632 TCC AAG TCC AGT GCA TTA GCT GCA GCT CGT AGA GCC GAA CGG GAG GGA GAA AAG AAG CAC 1691
413 S K S S A L A A A R R A E R E G E K K H 432

1692 AGA CAA AAC AAT GAA GCA GAG AGA CTT TTA TTG GCT GGA GTG GAG ATT CCA GAA CCT GTT 1751
433 R Q N N E A E R L L L A G V E I P E P V 452

1752 TTC TTC TGT ACC ATA GAA CCC CCA TCA CTG TCT AAG CAG CCA GAT TTG GAA CAT GCG TTG 1811
453 F F C T I E P P S L S K Q P D L E H A L 472

1812 AAA TGT CTT CAG CGT GAA GAT CCC AGT TTG AAA GTG AGG CTA GAT CCT GAC TCT GGA CAA 1871
473 K C L Q R E D P S L K V R L D P D S G Q 492

1872 GCA CCA CCC AGA AGC TGT CCA CAT CAA TTA TAT TCA TCT TCC ATT GGC CAG AAA TTT GGC 1931
493 A P P R S C P H Q L Y S S S I G Q K F G 512

1932 TTT GCG GTT GGG AAG TGT TGG TAT TTA GCT AAG TGT TGG CAT CAT TAG CCTTAGTCAG ATG 1992
513 F A V G K C W Y L A K C W H H * M 1

1993 CTA CTT GCC CAG CAG ATA ATT CAA CTA TCA AAG AAG AAG GAG AGA ATG GCC AGG TAA TTCA 2053
2 L L A Q Q I I Q L S K K K E R M A R * 20

2054 TTGTTTATTGATTCATTATTCCAGTCAGAAAAGTCAAAAGCA ATG TTT TGA TGTAAAAGCCCCGGGCCAAAAAGAA 2129
1 M F * 3

2130 TTACAAAGGAAACTGTAATTAATTAAAAATAAGT ATG TAT TAA CATCATACTTTTATATAGTATTCTGCAATAAAT 2205
1 M Y * 3

2206 ATAATTTAGCCTTAATCTTTTATTTCATTCCTCTGGTAACCAGACTGTTCTGTGTGGT ATG GGG GAG TTA CAT 2278
1 M G E L H 5

2279 ATA GAG ATT ATT CAT GAT CGA ATC AAG AGG GAA TAT GGA CTG GAG ACC TAT CTC GGG CCT 2338
6 I E I I H D R I K R E Y G L E T Y L G P 25

2339 CTC CAG GTG GCA TAT CGA GAG ACC ATC CTA AAC TCA GTT CGT GCC ACA GGT AAA AAG TAA 2398
26 L Q V A Y R E T I L N S V R A T G K K * 45

2399 AACATTTAGGATTTGCTTTTTCTCTCTTTCTGTCTTTCTTTCTCTATTTAGTTGCTGAAAGCAGTGTTCATAGTCACTA 2477

2478 ATG TCT TCT ATC AGA GGA GAC AGT TTT TAT GTA TGT GGG GTT TTT TTG TTT TTT TAG GATT 2538
1 M S S I R G D S F Y V C G V F L F F * 19

2539 TTGTTTTGAATACAGAAAGCCCTATTTGATTATCAGCCATAACATCTCCTGG ATG ACC ACT CAT CAA CCA AGT 2611
1 M T T H Q P S 7

2612 GTC CTT TTC AGT GGT CAT GGA ATA GAG TAA GCCACTGAAGTCG ATG GCT TGT CCC AGT CCT 2672
8 V L F S G H G I E * M A C P S P 6

2673 CAC TAT CAG AAC TAG TCTAAGAATTTGAAGTCA ATG TGT TAT TTT TAC TCT ATA AAG TAA AAGTA 2737
7 H Y Q N * M C Y F Y S I K * 9

2738 GAAAGTCAGTTTCAGTACTTAATAAAT ATG ATT ATT CTA TGC TGA TAGTAATCTTGAGAAACTATAGTTTTTA 2810
1 M I I L C * 6

2811 TACTTCTATTCCAA ATG TAA AAACTCACAATTTTTCTGCAT ATG ATT TTA ATT TGG TTC CCT TAA GC 2878
1 M * M I L I W F P * 8

2879 CAAAAAAC ATG ACA ACA ATT ATT GCT ACT CAC CCC CAA GAA GAA AGT GTG GGA AGG TTT GTG 2940
1 M T T I I A T H P Q E E S V G R F V 18

2941 GGA AAT TGT TTA AGA AAA AAA AAA AAA 2967
19 G N C L R K K K K 27

FIGURE 9 continued